DOCUMENT RESUME

ED 290 136 CS 009 029

AUTHOR Lumpkin, Donavon, Ed.; And Others

TITLE Evaluation in Reading: Learning, Teaching, Administering. Sixth Yearbook of the American Reading Forum.

SPONS AGENCY American Reading Forum.

PUB DATE 86

NOTE 369p.; For related documents, see CS 009 027-028.

PUB TYPE Viewpoints (120) -- Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC15 Plus Postage.

DESCRIPTORS Administrator Evaluation; *Administrator Role; Elementary Secondary Education; *Evaluation Criteria; Higher Education; *Reading Diagnosis; *Reading Instruction; Reading Research; Teacher Effectiveness; *Teacher Evaluation; *Teacher Role

ABSTRACT

Centering around the common theme of reading evaluation, the papers in this yearbook were either presented at or are based on presentations from the 1985 American Reading Forum conference. Titles of some of the papers and their authors are as follows: "Teacher Competency Testing" (S. Kossack); "Principals as Instructional Leaders" (J. Laffey); "Issues Related to Testing" (W. Otto); "Issues Related to Labels" (K. Camperell); "Issues Related to Personal Context" (B. Hays); "A View of a Statewide Evaluation" (J. Handler and D. Carlson); "The Role of the Reading Educator in the Training of Elementary School Principals" (R. Kurth); "Are Student Interviews a Legitimate Evaluation Tool for Assessing School Reading Programs?" (J. Cassidy and others); "Evaluation of Adult Illiterates' Oral Language Associations with the Paradigmatic/Syntagmatic Inventory" (J. Dinnan and L. Pulling); "Nonliterate Adults' Knowledge of and Attitudes toward Microcomputers for Reading Instruction" (L. Wachter and E. Askov); "Application of Teacher Effectiveness: Research" (T. Blair and J. Young with a reaction by C. Stice); "Reading Evaluation in the Classroom" (L. Gentile); "An Analysis of Linguistic and Comprehension Strategies of Low Ability Readers" (V. Risko); "College Reading Programs 1958-1978" (A. Mallery); "Developmental Reading Program Innovations and Practices at the University of Cincinnati" (T. Bullock and D. Madden); "Guidelines and Practices of Evaluation of State Reading Programs" (P. Smith and S. Rebottini); "Teacher Perceptions of Microcomputer Courses" (P. Ransom); "Evaluation of Variations among Reading Improvement Students through Three Decades" (M. Harshbarger); "A Comparison of Parent and Teacher Perceptions of Student Behaviors" (J. Williams); "Evaluating Perceptions of Learners at Different Levels" (D. Lumpkin); "Research in Teaching Context Clues" (J. Elliott); "Reading Inventories of the Future" (W. Henk with a reaction by L. Tomlinson); "The Cloze Reading Inventory as a Qualitative and Quantitative Measure of Reading Ability" (R. DeSanti with a reaction by C. Stice); "Content Reading" (R. Telfer); "Preliminary Development of a Screening Instrument for Learning Disabilities" (C. Cloer, Jr.); "The Mandate: Evaluation of Teacher Candidates by Use of Competency Tests" (R. Flippo); "How Effective Are Quantitative Teacher Evaluations?" (B. Herrmann); "Evaluation of a Reading Program" (B. Wise and J. Shaver); "Students' Perceptions of Their Study Habits" (B. Townsend); "Evaluation of

Learning Disabled College Students" (K. Norlander and others); "Reading and Learning: The Learning Abilities of Three Types of Poor Readers" (R. Bloomer and others with a reaction by R. Flippo); "Ethnographic Research on Junior Great Books" (N. Boraks and others); "Evaluating Materials" (M. Collins); "Evaluation of Teaching with Discussion" (D. Alvermann); "New Ways to Use Student Writing Samples in Reading Diagnosis" (P. Duncan). References for each article are included. (SKC)

****
*   Reproductions supplied by EDRS are the best that can be made   *
*   from the original document.   *
****

ED290136

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Peggy E. Ranson
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS009029

EVALUATION IN READING:
LEARNING
TEACHING
ADMINISTERING

SIXTH YEARBOOK OF THE AMERICAN READING FORUM

1986

Edited by
DONAVON LUMPKIN
MARY HARSHBERGER
PEGGY RANSOM

CS 009029

BALL STATE UNIVERSITY
MUNCIE, INDIANA

The ideas expressed in the papers published
in the Yearbook are those of the authors and
do not necessarily represent the viewpoints
of the American Reading Forum

Copyright 1986
All rights reserved. This book or parts thereof
may not be reproduced in any form without
permission of the publishers.

THE AMERICAN READING FORUM
Memberships:
4900 Ocean Blvd. #502
Sarasota, Florida 33581
Subscriptions:
Ball State University
T.C. 307C
Muncie, Indiana 47306

PRINTED IN THE UNITED STATES OF AMERICA
By
BALL STATE UNIVERSITY
Muncie, Indiana 47306

AMERICAN READING FORUM CONFERENCE

Evaluation in Reading: Learning—Teaching—Administering

December 12, 13, 14, 1985
Sarsasota, Florida

Board of Directors

1987
Martha Collins
Charles Peters
Peggy Ransom

1986
Eunice Askov
Bernard Hayes, Vice Chair

1985
Byron Calloway
Marian Tonjes, Chair

Secretary to the Board: Sharon Kossack
Treasurer to the Board: Thomas Cloer

Conference Proposal Review Committee

Marino Alvarez
Donna Alverman
Nancy Boraks
Roger DeSanti
Beth Herrmann
Robert Jennings
Bob Jerrolds
Chester Laine
Anne Mallery
George Mason
Wayne Otto
Norman Stahl
Sarah Weidler
Deborah Evans

Program Co-Chairman

Brenda Kolker—Paul Terwilliger

Forum Committee

Norman Stahl, Chairman
Eunice Askov
Tony Manzo

Membership/Registration Committee

Ann Hall, Chairman
Betty Yarborough
Ula Casale
John Wallenberg

Local Arrangements Committee

Evelyn Spache, Chairman
Betty Raygor
Glenda Lofton

The *Sixth Yearbook* of the American Reading Forum is composed of papers accepted for publication by the Editorial Board from those submitted by presentors at forums, problems courts and sessions of the 1985 Annual Conference. The Board of Directors and members of the planning committees for that conference contributed significantly to this yearbook.

EDITORIAL ADVISORY BOARD

Donna Alvermann	University of Georgia
Kay Camperell	Utah State University
Lucy Dahlberg	Governors State University/Illinois
Gerald Duffy	Michigan State University
Joan Elliot	Indiana University at Pennsylvania
Rona Flippo	Fitchburg State College/Massachusetts
Mary Harshbarger	Ball State University
Robert Jennings	Fort Hays State University
Sharon Kossack	Florida International University Bay Vista Campus
Ruth Kurth	North Texas State University
Chester Laine	University of Cincinnati
Donavon Lumpkin	Ball State University
Anthony Manzo	University of Missouri/Kansas City
George Mason	University of Georgia
George McNinch	West Georgia College
Donna Norton	Texas A & M
Wayne Otto	University of Wisconsin/Madison
Peggy E. Ransom	Ball State University
Evelyn Searls	University of South Florida
Norman Stahl	Georgia State University
Joan Williams	Ball State University

The editors of EVALUATION IN READING: LEARNING-TEACHING-ADMINISTERING, the sixth yearbook of the AMERICAN READING FORUM, wish to express genuine appreciation to members of the 1986 Editorial Board. These members have served as advisers, reviewers, and supporters in the publication of this yearbook. Since its beginning the yearbooks of the AMERICAN READING FORUM have been refereed publications of papers presented at annual conferences. Manuscripts based on conference presentations are reviewed by the Editorial Advisory Board before acceptance by the editors for inclusion.

Don Lumpkin
Mary Harshbarger
Peggy Ransom

Foreword

The papers in this yearbook have been organized into groupings based on the format in which they were presented at the 1985 Annual Conference in Sarasota, Florida. The keynote speeches from the first General Session and the presentations at Problem Courts are clustered together in the initial section. These articles often employ the less formal first person in order to retain the quality of give and take interchange which characterize the consideration of problems, issues and concerns by the American Reading Forum.

Articles prepared from papers presented at other sessions and meetings have been grouped in the second section of the yearbook. Each article contributes in some way to a broad perspective of Evaluation in Reading: Learning - Teaching - Administering. This theme selected for the Annual Conference focused attention on an area of major factors exercising current impact on reading and in the education of reading teachers. In a decade dominated by concerns related to literacy and functional reading competencies the need for thoughtful scrutiny of evaluation in reading is obvious. This evaluation should encompass not only performance of individual readers but also the effectiveness of teachers and administrative personnel who influence this area of development.

The editors of the yearbook and members of the American Reading Forum hope that publication of these proceedings will contribute to a more complete understanding of the concerns and challenges related to reading education today. To assist in achieving this objective the editors and reviewers have recommended inclusion of a wide array of articles based on the papers presented. Some examine aspects of reading previously subjected to extensive study, reviewing status and expanding perceptions toward added dimensions. Others explore new possibilities, expanding horizons that can lead toward the future. If consideration of topics from this broad spectrum stimulates efforts which promote improvement and growth in literacy and reading, our goals will have been accomplished.

D. L.
M. H.
P. R.
Muncie, Indiana
1986

TABLE OF CONTENTS

Keynote Addresses and Problems Courts

OVERVIEW OF EVALUATION AND TEACHING

TEACHER COMPETENCY TESTING: A DOUBLE-EDGED SWORD 1
Sharon Kossack

OVERVIEW OF EVALUATION OF ADMINISTRATORS

PRINCIPALS AS INSTRUCTIONAL LEADERS: NOW IS THE TIME! 17
James Laffey

DIAGNOSIS OF READING PROBLEMS: WHAT ARE THE ISSUES?

ISSUES RELATED TO TESTING 28
Wayne Otto

ISSUES RELATED TO LABELS 30
Kaybeth Camperell

ISSUES RELATED TO PERSONAL CONTEXT. 35
Bernie Hays

REACTION PAPERS
Commentary on Otto, Camperell & Hays 38
Ruth J. Kurth

Commentary on Otto, Camperell & Hays 41
Carol Dana

A VIEW OF A STATEWIDE EVALUATION: IMPLICATIONS OF TENNESSEE'S CAREER LADDER 44
Janet Handler & Deborah L. Carlson

THE ROLE OF THE READING EDUCATOR IN THE TRAINING OF ELEMENTARY SCHOOL PRINCIPALS 53
Ruth J. Kurth

ARE STUDENT INTERVIEWS A LIGITIMATE EVALUATION TOOL FOR ASSESSING SCHOOL READING PROGRAMS? 59
Judith K. Cassidy, Lawrence G. Erickson, Kathleen Gaare-Wiese & Deborah J. Rogers

EVALUATION OF ADULT ILLITERATES ORAL LANGUAGE ASSOCIATIONS WITH THE PARADIGMATIC/SYNTAGNATIC INVENTORY. 68
James A. Dinnan & L. Vern Pulling

Presentations at Sessions

NONLITERATE ADULTS' KNOWLEDGE OF AND ATTITUDES TOWARD MICROCOMPUTERS FOR READING INSTRUCTION 74
Louise D. N. Wachter & Eunice N. Askov

THE DEVELOPMENT OF AN INSTRUMENT TO MEASURE TEACHER ATTITUDES TOWARD THE COMPUTER 82
Sidney E. Benton, Judy Long & Bob W. Jerrolds

APPLICATION OF TEACHER EFFECTIVENESS: RESEARCH FINDINGS 92
Timothy R. Blair & June Thomas Young

REACTION PAPER
Commentary on Blair & Young 98
Carole F. Stice

READING EVALUATION IN THE CLASSROOM: BROADENING ASSESSMENT AND INTERVENTION 99
Lance M. Gentile

AN ANALYSIS OF LINQUISTIC AND COMPREHENSION STRATEGIES OF LOW ABILITY READERS 105
Victoria J. Risko

COLLEGE READING PROGRAMS 1958-1978 113
Anne L. Mallery

DEVELOPMENTAL READING PROGRAM INNOVATIONS AND PRACTICES AT THE UNIVERSITY OF CINCINNATI. 126
Terry Bullock & Deborah Rosencrans Madden

GUIDELINES AND PRACTICES OF EVALUATION OF STATE READING PROGRAMS: HOW VIABLE ARE THE CRITERIA?. . . . .
Patricia K. Smith & Sandra L. Rebottini

TEACHER PERCEPTIONS OF MICROCOMPUTER COURSES. 143
Peggy E. Ransom

EVALUATION OF VARIATIONS AMONG READING IMPROVEMENT STUDENTS THROUGH THREE DECADES 149
Mary Harshbarger

A COMPARISON OF PARENT AND TEACHER PERCEPTIONS OF STUDENT BEHAVIORS . 156
Joan Williams

EVALUATING PERCEPTIONS OF LEARNERS AT DIFFERENT LEVELS. 167
Donavon Lumpkin

RESEARCH IN TEACHING CONTEXT CLUES: AN ANALYSIS OF FOUR BASAL READING SERIES AND READING METHODOLOGY TEXTBOOKS. 174
Joan Elliott

READING INVENTORIES OF THE FUTURE 186
William Henk

REACTION PAPER: COMMENTARY ON HENK 192
Louise Tomlinson

THE CLOZE READING INVENTORY AS A QUALITATIVE AND QUANTITATIVE MEASURE OF READING ABILITY 194
Roger J. DeSanti, Renee M. Casbergue & Vicki G. Sullivan

REACTION PAPER: COMMENTARY ON DESANTI, CASBERGUE & SULLIVAN . 204
Carole F. Stice

CONTENT READING: PREACHING AND PRACTICING 205
Richard Telfer

PRELIMINARY DEVELOPMENT OF A SCREENING INSTRUMENT FOR LEARNING DISABILITIES, GRADES 2-8. 212
Carl Thomas Cloer, Jr.

EFFECTIVENESS OF TEACHER OBSERVATION IN DETECTING VISUAL DIFFICULTIES 222
John Alvarez & Sharon Kossack

THE MANDATE: EVALUATION OF TEACHER CANDIDATES BY USE OF COMPETENCY TESTS 229
Rona F. Flippo

HOW EFFECTIVE ARE QUANTITATIVE TEACHER EVALUATIONS? 233
Beth Ann Herrmann

AN INVESTIGATION OF THE RELATIONSHIPS BETWEEN STUDENT ACHIEVEMENT AND TEACHERS PERCEPTIONS OF INSTRUCTIONAL HINDRANCES IN THE ELEMENTARY SCHOOL 240
Donald D. Neville, Katherine Richmond & Evelyn F. Searls

REACTION PAPER: COMMENTARY ON NEVILLE, RICHMOND AND SEARLS. . 251
Marino Alvarez

EVALUATION OF A READING PROGRAM: TEACHER JUDGMENT AND PUPIL PLACEMENT 253
Beth S. Wise and Judy C. Shaver

STUDENTS' PERCEPTIONS OF THEIR STUDY HABITS 261
Brenda S. Townsend

EVALUATION OF LEARNING DISABLED COLLEGE STUDENTS: A PROFILE ANALYSIS APPROACH. 265
Kay Norlander, Anthony Paolitto & Ania Czajkowski

A COMPARISON OF AUDIO-LINGUAL-HAPTIC AND INTERACTIVE VOCABULARY APPROACHES FOR COMPREHENSION TRANSFER OF SECONDARY ESL STUDENTS . . 278
Virginia Laudermilk and Sharon W. Kossack

REACTION PAPER: COMMENTARY ON LAUDERMILK AND KOSSACK. 283

READING AND LEARNING: THE LEARNING ABILITIES OF THREE TYPES OF POOR READERS 286
Richard Bloomer, Michael Martin & Elizabeth Levesque

REACTION PAPER: COMMENTARY ON BLOOMER, MARTIN AND LEVESQUE . . 296

THE APPLICATION OF SINGLE-SUBJECT MULTIPLE - BASELINE DESIGNS TO EDUCATION 297
Sarah D. Weidler and Connie M. Maclay

REACTION PAPER: COMMENTARY ON WEIDLER AND MACLAY 305
Carole F. Stice

ETHNOGRAPHIC RESEARCH ON JUNIOR GREAT BOOKS: DISCUSSION LEADER BEHAVIOR 306
Nancy Boraks, Julie Early & Eileen Sable

EVALUATING MATERIALS: DO THEY TEACH PHONICS GENERALIZATIONS? . . . 317
Martha D. Collins

EVALUATION OF TEACHING WITH DISCUSSION: A TRIANGULATED INQUIRY . . 329
Donna E. Alvermann

THE ROLE OF SELECTED PRINCIPALS IN THE EVALUATION OF THE TEACHING OF READING 337
Rebecca M. Moore, Sally M. Boyette & Sylvia McCoy Carter

NEW WAYS TO USE STUDENT WRITING SAMPLES IN READING DIAGNOSIS 344
Patricia H. Duncan

TEACHER COMPETENCY TESTING: A Double-Edged Sword

SHARON W. KOSSACK
Florida International University

Teacher competency testing, a parallel to the professional tests in medical and professional fields, is on an unprecedented rise and offers the field a double-edge blade. It has the potential of offering STANDARDS, DEFINITION, and CONTENT.

STANDARDS. Teacher testing has the potential of improving the quality of education by assuring higher-quality candidates for teaching positions; or the standards can be lowered to a minimum, making this level our goal.

DEFINITION. The qualities of effective teachers can at long last be defined via research surrounding teacher testing as tests are shown consistent with job qualifications, a legal requirement for a test to be used to disallow "incompetent" candidates; or we may discard as victims those candidates who fail to meet arbitrary minimum cutoff points on arbitrary tests which are not consistent with job responsibilities.

CONTENT. Will tests be developed, for they certainly do not exist at the present time, which will establish essential aspects of teacher effectiveness; or will tests be chosen haphazardly whose content is unrelated to the act and art of teaching?

PRESENT STATUS/PROGRESS OF TEACHER TESTING

In 1980, eight Southern states (16% of the nation) required teacher competency testing; five years later that figure has tripled, with 27 states (32% of the nation) now requiring teacher examinations. (See Table I) The increase, alone, is not of concern, for if teacher testing has the potential of offering the profession standards, content, and definition, this increase is to be expected. Of greater concern is the source and speed of implementation.

Source. Over half of the teacher competency movements have been legislatively mandated (52%). (See Table I) The other cluster has been initiated by State Boards/Departments of Education. It is important to realize that these efforts have NOT been initiated by teacher organizations or teacher training programs intent on governing and improving their own ranks; indeed, there is often little input from them. With the major initiation from legislators, there should be concern that the choice of testing and tests will be lay-advised and inappropriate for the teaching environment.

TABLE I

TEST-COMPETENCY IMPLEMENTATION BY STATE

	**Initiator	Year of Initiation	Year of Implementation	***Test	****Skill Focus			When Test is *****Required
Alabama	S	1980	1981	S		P	C	X
Arizona	L	1980	1980	S	B	P	C	X
*Arkansas	L	1979	1983	N	B	P	C	X
California	L	1981	1982	S	B		C	
Colorado	L	1981	1982	S				N
Connecticut	S	1982	1985	S				
Delaware	S	1982	1983	N				
*Florida	L	1978	1980	S	B	P		N/X/I
*Georgia	S	1979	1979	S			C	X
Indiana	L	1984	1985	N				
Kansas	L	1984	1986	S				
Kentucky	L	1984	1985	N				
*Louisiana	L	1979	1979	N			C	
Maine	L	1984	1988	N				
*Mississippi	S	1975	1977	N	B	P	C	
Nebraska	L	1984	1984					
New Hampshire	S	1984	1985	N				
New Mexico	S	1981	1983	N				
New York	S	1980	1984	N	B	P	C	X/I
*No. Carolina	S	1979	1981	N	B	P		
Oklahoma	L	1980	1982	S				
*So. Carolina	L	1979	1982	N/S			C	X
Tennessee	S	1980	1981	N			C	N
Texas	S/L	1981	1986	S	B		C	
Virginia	L	1979	1980	N	B	P	C	X
Washington	S	1984	TBA	N	B	P	C	
West Virginia	S	1982	1985	S				

Source: "Testing for Teacher Certification," Issuegram #7, January 3, 1983 (ED 231 814), Education Commission of the States; 1860 Lincoln St., Suite 300; Denver, CO 80295

*States which initiated teacher testing prior to 1980.
**Initiator: S=State Board/Department of Education; L=Legislature
***Test Source: S=State Developed; N=Nationally Normed (NTE/ACT)
****Skill Focus: B=Basic Skills (Reading, Mathematics, Writing); P=Professional Practices; C=Content Knowledge
*****When Required: N=Entrance to Program; X=Exit from Program; I=Internship

Rapidity. Of the 27 states who have sought implementation of teacher testings, 11 have required it within one year of initiation. (See Table I) One would have to question how such massive testing could be thoughtfully planned, organized, and implemented in the context of a calendar year; which considering the mismatch in academic and legislative years would be less than 6 months.

Teacher trainers have lost the initiative for self-regulation. No longer do teachers have the option of policing.

WHY TEACHER COMPETENCY TESTING?

Is there a need for teacher competency testing? What are the reasons there is testing now? Answers to these questions follow.

Student Achievement Declines. The public has observed a 15 year decline in SAT test scores, widely-publicized legal cases (e.g. Lau vs. Nichols) emphasize the incidence of illiterate high school graduates. Recent Gallup Polls have indicated the loss of public support for education, as 85-95% of the respondents favored teacher testing (Cole, 1979; Nassif, 1982).

Education Majors. Recent studies have indicated that education majors have SAT test scores that are third from the bottom and declining. Results from the teacher testing programs that were instituted in the early 1980's indicate a 20% failure rate on basic skills subtests of the examinations. Open admission has become open exit as teacher trainers fail to meet the obligation of "gatekeepers" to prevent the incompetent from obtaining certifiable degrees.

Inservice Teacher Performance. Superintendents report 5-15% of practicing teachers give inadequate job performances; and this is considered an understatement (Legislative Review 1980). The number of teachers dismissed has gradually increased (1975--684; 1976--766; 1977--847; American School Boards Association Survey, 1977).

Political Opportunity. Politicians are quick to sense an undercurrent of dissatisfaction that would work to their political advantage. Any politician that could pass legislation to improve education has the potential of reaching many constituents. Media has added to the controversy, emphasizing the crisis in education in numerous articles (e.g., "Failing Teachers," *The Miami Herald*, 1982; "Help! Teacher Can't Teach," *Time*, 1980). At a national political seminar, 63 legislators from 38 states identified "competency of teachers" as one of their top five educational priorities along with funding, accountability, equity of state aid allocations, and special education (*Legislative Review*, 1980). Note that of the 27 states which require teacher testing for certification, over half were legislatively initiated, further testament that this is fast becoming a political issue. This has reached the national arena as governmental guidelines have been devised which facilitate teacher assessment, specifying: (1) "fair and substantial relationship" to governmental objectives for education, (2) nondiscriminatory intent, (3) validation in terms of job performance, and (4) business necessity test under Title VII of Civil Rights Act of 1964. There have been multiple challenges of teacher testing, despite its 6-year tenure.

WHAT TESTS ARE BEING USED?

Should a cluster of educators be polled, they would hardly be expected to agree on the identification of a specific existing test appropriate for teacher competency testing, nor would they be able to agree on the content of one to be devised. They would be expected to agree on one fact: there is no valid test of teacher competence currently in existence, especially of the pencil-and-paper variety, that would meet the governmental guidelines cited above. There is general agreement that we have not yet reached the level of technology in measurement to tap essential teacher skills (Pugach and Raths, 1983). Yet, 27 states are currently testing teacher competence. What tests are being used?

Of the 27 states using teacher testing, over half have chosen to use national, normed tests (most often the National Teacher Exam). The others use state-developed tests or a combination of national and state tests. There are a number of difficulties with use of these tests.

Student Learning. There is little support to indicate there is a positive relationship between teachers' NTE scores and the achievement of students (Darling-Hammond and Wise, 1983; Lanier, 1981; Quirk, et al, 1973; Weinberg, 1973; Soar, 1981; Summer and Wolfe, 1975).

Predictive Validity. The Austin School District abandoned teacher testing because there are only a "small fraction of teacher competencies that could be adequately measured by tests, the large cost of developing valid and legally defensible measures, and the threatened negative impact on district recruitment and hiring efforts" (Nassif, 1982; Holley, 1980). There is "no set of teacher competencies that have been empirically validated by research as predictive of success in teaching or essential for effective learning to take place" (Schalock, 1980).

Content Validity. There is no agreement among educators on what a public school teacher should know and be tested on in order to qualify for a certificate (Hyman, 1984). Worse, cutoff points have been established in haphazard and arbitrary ways. Cutoff points, identifying the competent from the incompetent, have been data free--we do not know that those failing to meet the cutoff point would NOT make excellent teachers, as we do not know that those failing the Bar Exam or medical test would make poor practitioners.

Basic Skills. This domain offers the greatest defensible use for teacher testing, in that it would serve to weed out those candidates with demonstrable basic skill deficiencies, using the logic that a teacher cannot teach with a lack of basic communication and computational skills to convey and measure ideas. Additionally, some limited research has shown that verbal skills of teachers have been shown to be more closely related to student achievement than other existing factors (Seeley, 1979).

Content Subject-Area Information. The face validity of content area testing is convincing on a logical level: teachers cannot teach content they do not, themselves, know. However, what is the "content" for elementary educators, early childhood teachers, special education professional? These "content tests" generally reduce to a mix of methodology, educational philosophy rather than subject matter. However, knowledge of content does not guarantee competent teaching. Seeley (1979) underscores that fact with a quote from a superintendent:

> ...incompetent teachers divide into two groups: "Those who don't have it" (who lack teaching skills) and "those who don't use it" who are able, but don't give adequate effort).

The danger of subject-area tests involve those teachers who are capable of doing a credible job, but who chose not to do so, or lack other, unmeasurable, teaching skills.

Professional Knowledge/Skills. There is no evidence submitted to show that the tests we have available to use validly measure a teacher's knowledge and skills. The NTE indicates it tests the prospective teacher's academic preparations (NTE Policy Council, 1979). NTE directly states that there is only a relationship between the tests and "the content of teacher education programs" (1983-4 Bulletin of Information). They do not claim that the tests measure skills needed in the actual performance of teaching (Hyman, 1984).

Thus, knowledge of philosophers (e.g., Dewey), psychologists (e.g. Skinner, Piaget), and the principles/ethics of education and best teaching practices are often included in teacher testing (48% of the states). This threatens to lend preference for particular pet ideologies of test constructors, constrain teacher preparation programs to teach to the test content rather than present a wide variety of approaches and make application to a number of classroom situations.

General Knowledge. The General Education subtests of the NTE tests candidates on knowledge of composers, articles, authors, scientific principles, and history--are these related to the act of teaching? Quirk (1973) found general knowledge scores to correlate moderately with college GPA, but to be poor predictors of ratings assigned to student in clinical experiences (median correlation of .05) and of subsequent assessments of teaching effectiveness evaluated by principals and supervisors on the job (median correlation of .10).

While general knowledge might be a desirable attribute, it cannot be shown to be essential to job performance (Soar, 1981), and discriminates against minorities (McDaniel, 1977; Ivie, 1982), thus violating governmental guidelines.

The violation of minority rights deserves more consideration, in light of projected minority teachers compared to a student population which is expected to double by 1990. Some school systems face a minority student population of 30-50%. How is this extreme depletion of minority teachers going to affect the performance of minority students who face a non-minority teaching staff for their entire learning career; where are their role models? This is not to imply the tests overtly discriminate against minorities, but the arbitrary and unvalidated cutoff points have been established at the point of greatest discrimination against minority candidates (Smith, 1984). (See Table II, III) In Georgia, 34% of the Blacks pass, as opposed to 87% of the Caucasian candidates (Hansen, 1983). This is mirrored in Louisiana (1983), with 15% Black pass rate, 78% Caucasian pass rate; and in Florida (Andrews, 1983) with 35% Black, 90% Caucasian, 51% Hispanic, and 81% Asian/Indian pass rates; California (Savage, 1983) with 26% Black, 67% Caucasian, and 38% Hispanic pass rates; and Texas (CBEST, 1984) with 4% Black and 16% Hispanic pass rates. It is significant to note that "a significant number of states that have adopted and have already implemented competency assessment of teachers are also states that are under federal mandate to desegregate their institutions of higher education" (Smith, 1984).

Oversight Domains. Essential teacher competencies not tested by the above measures are:

> ...teacher who have the temperament, patience, compatibility both to get along with young people and plan effective teaching (Seeley, 1979).
>
> ...emotional stability, eliminating those whose problems may be rooted in the home or professional areas.
>
> ...flexibility, eliminating the veterans who cannot keep up, who may be experiencing a late career slump or who have lost their drive, interest, and adaptability, or who have trouble handling newer instructional programs or different student bodies (Staff Dismissal, 1978).
>
> ...teacher aptitude, interests, attitudes, motivation, maturity, creativity, and other social characteristics (Harris, 1981; Hyman, 1984).

TABLE II

PERCENTAGE OF MINORITY TEACHERS: NATIONAL

	1980	1990 (projected)
Black	8.6%	
Hispanic	1.8%	
Asian/Indian	< 1.0%	
Overall	2.5%	< 5.0%*

TABLE III

RESULTS OF STATE TESTING: MINORITY JEOPARDY

	Black	White	Hispanic	Asian/Indian
Georgia (Hansen, 1983) State test	34%	87%		
Louisiana (Teach test, 1983) NTE	15%	78%		
Florida (Andrews, 1983) State test	35%	90%	51%	81%
California (Savage, 1983) CBEST	26%	67%	38%	
Texas (CBEST, 1984) P-PST	4%		16%	

NORMATIVE VERSUS CRITERION-REFERENCED TESTS

Norm-referenced tests provide no corrective feedback for candidates or teacher training programs. This becomes serious as some states link teacher education program approval status to student test performance. Thirty-eight teacher education programs from 18 different colleges lost program approval because 80% of their graduates failed to pass each of the three parts of the State Teacher Examination in Florida (In Brief, 1983; Smith, 1984). Those programs who lost approval do not know how to alter teacher preparation programs to alleviate this problem.

WHEN ARE TESTS ADMINISTERED

Ethical issues arise when one considers when the evaluations are administered. Professionals who have invested time and money in undergraduate and graduate programs of study should not be gauged incompetent at their completion. What are the national trends in this area (Palaich, 1985; Sandefur,1983)?

Preservice. (31 states)

(a) Entrance Requirements (28 states)
(b) Exit Requirements (24 states)

Induction (37 states)

(a) Initial Certification (30 states)
    Basic Skills (27 of 36 states)
    Professional Practices (21 of 36 states)
    Subject Area Content Knowledge (20 of 36 states)

(b) First Year Approvals Pending Performance Evaluations (19 states)

Inservice (38 states)

(a) Recertification Requirement
    Staff Development Post-tests
    Standardized Tests (NTE/ACT/SAT/CAT) 17 states
    Criterion Referenced Tests (16 states)
    Performance Evaluations (20 states)

Palaich (1985) concludes that 31 (61%) of the states have shown significant preservice activities, 37 states (72%) have shown significant induction activity, and 38 states (74%) have shown significant inservice activity. Only two states have remained uninvolved in the teacher testing movement. In 50% of the states, there was significant activity in all three, suggesting most states are taking a comprehensive focus.

Virtually no state showed high activity in only two areas. If a state tended to concentrate in only one area, it was usually inservice (8 states) or induction (6 states) and tend to involve innovative policies (e.g., internships for beginning teachers, career ladders). Many states are taking an incentive versus an assessment view (Palaich, 1985).

IMPACT ON PUBLIC OPINION

Publicity (e.g., "Help!", Times, 1980) about declining enrollments in teacher education programs and the attendant lowering of SAT scores of teacher candidates (Weaver, 1978) have increased public worries about education; but public opinion is not as negative toward public schools and teaching as the portrayals offered by mass media suggest.

Weiler (1982) indicated that participants in the Gallup Polls have assigned progressively more "bad" (C-F) than "good" (A-B) overall grades to schools from 1976-81. Gallups' 1981 analysis reminds us that residents consistently tend to give their local community schools higher ratings than the public schools nationally, Weiler suggests that the apparent furor over the quality of public schools lies with a perception of national problems; such lack of confidence does not appear to be reflected in public satisfaction with local schools. The difficulty of hiring good teachers ranked only 5th in the major problems facing the schools; lower than funding (Gallup, 1981). By 1980, "hiring good teachers" was ranked 7th, a lessening concern.

Palaich (1985) feels that national teacher testing will improve the public attitude toward the schools:

> Public attitudes toward education seem to improve when regulations become more rigorous--when entrance, certification, and recertification requirements are strengthened.

Pugach and Raths (1983) suggest that public opinion would be modified only if three criterion were met: (1) tests are shown to test the essential competencies of teaching, (2) both preservice and inservice teachers were required to take them, and (3) the tests must be difficult to pass.

CONCLUSIONS AND RECOMMENDATIONS

From the review of teacher competency testing the following conclusions and recommendations could be made.

1. Teacher testing should be only one aspect of a broader, coordinated improvement effort which includes all points in the contimuum of professional preparation (Pugach, 1983).

> When such testing is adopted in a short time on a large-scale fashion, ... there may be a tendency to overlook the potential reactions of both the school and the general public as well as a failure to anticipate misuses of the results ... (which) might work to undermine the very real contributions that teacher testing could make to the overall efforts to improve the quality of the teaching force (NAEP, 1982a).

2. Raise admissions standards to Teacher Education programs.

Scholastic Aptitude Test scores of education majors are lower than those of other professional majors; to raise the quality of candidates, some states (33%) are requiring entrance performance criteria. Pugach (1983) suggests this should be limited to basic skills assessment and higher GPA requirements at the time of admission to teacher training. This provides hard data to faculty who can apply their "gatekeeping" responsibilities to suggesting remediation or redirection to other fields. Technical (validity), ethnic (minority, uncertifiable candidates after a four-year investment of time and money for training), and economic factors preclude the application of other, possibly more valid, domains.

Potential consequences might involve exacerbation of the existing teacher shortage and discrimination against minority candidates.

3. Strengthen the institutional accreditation process.

Program faculty should screen, prepare, counsel, evaluate teachers during preservice training during course contact, advising, field site observations, as a means of addressing all aspects of the professional development of a teacher. Pugach (1983) calls for the institution of explicity and rigorous teacher evaluation systems within each teacher education program.

Consequences involve complex, time-consuming processes not without legal and professional implications.

4. Alternatives to testing

Some states (16%) are requiring new teachers to serve a one-year paid internship in the classroom before receiving full certification, allowing for evaluation, feedback, and training experience. Additional on-the-job training, through graduate coursework or workshop can upgrade teacher skills.

Consequences--quality of supervision and evaluation are critical and could undermine effectiveness; mandatory workshops result in reluctant learners; quality and evaluation of workshops can undermine effort.

5. Reform Teacher Education

New demands being placed on teachers (new knowledge, diverse populations, social problems) require a total overhaul of teacher education curriculum. There is a strident and continuing call for more content in lieu of the preponderance of methods courses (include more writing and mathematics courses to round out students' basic skills).

Consequences--there is a real need for consensus of content in teacher training. Perhaps it is wishful thinking, but teacher educators could consider this an auspicious time to band together to work out differences and do cooperative investigations to carve a new tomorrow in teacher education.

6. Challenge tests

Individual areas of specialization are not yet targeted, except in isolated states (Georgia, e.g., gives a reading teacher examination). Reading educators need not disenfranchise themselves with regard to future testing of reading specialists. They can place themselves, now, in a proactive role by cooperating in identifying the primary role performances required of various levels of reading professionals and drafting objectives and possibly test items to coordinate with same. Were this to be done on a national scale, through a teacher training organization, the impact could be notably innovative.

Educators need to be the leaders in the teacher competency scene. They should be included in the use or non usage of tests, the writing of tests, and the education of communities to the position of competency testing. Educators should have an impact in this area.

TEACHER TESTING BIBLIOGRAPHY

American Psychological Association. (1974). American Educational Association, and National Council on Measurement in Education, Standards for Educational and Psychological Tests and Manuals, Washington, D.C.: American Psychological Association.

American School Board Association. (1978). Staff Dismissal: Problems and Solutions.

Andrews, T.E. (1983). Teacher Education Reports.

Berk, R. (1980). Criterion Referenced Measurement: The State of the Art, Baltimore: Johns Hopkins University Press.

Civil Rights Act of 1964, 42 U.S.C.A., 2000.

Cole, R.W. (1979). Minimum Competency Tests for Teachers: Confusion Compounded, Phi Delta Kappan, 61, 233.

Darling-Hammonds, L. and Wise, A.E. (1983). Teaching Standards or Standardized Teaching? Educational Leadership, 41, 66-69.

Educational Testing Service. (1983). National Teacher Examinations Technical Handbook. Princeton, N.J.: Educational Testing Service.

Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor & Justice. (1978). Adoption of Four Agencies of Uniform Guidelines on Employer Selection Procedures, Federal Register, 43, 38290-38315.

Flippo, R. and Foster, C. (1984). Teacher Competency Testing and Its Impact on Educators. Journal of Teacher Education, 35, 10-13.

Gallup, G.H. (1979). The 11th annual Gallup Poll of the public's attitudes toward the public schools. Phi Delta Kappan, 61, 33-45.

Gallup, G.H. (1979). The 12th annual Gallup Poll of the public's attitudes toward the public schools. Phi Delta Kappan, 62, 33-48.

Gallup, G.H. (1981). The 13th annual Gallup Poll of the public's attitudes toward the public schools. Phi Delta Kappan, 63, 33-47.

Harris, A. (1983). Competency Tests Taking a Toll on Black Teachers, Baton Rouge Sunday Advocate, 7.

Hazard, W.R., Freeman, L.D., Eisdorfer, S., and Tractenberg, P. (1977).

Legal Issues in Teacher Preparation and Certification, ERIC Clearinghouse on Teacher Education: Washington, D.C.

Holley, F. F. (1980). An Alternative Approach to Assuring Teacher Competence, A paper presented at American Educational Association.

Hoover, M. (1984). Teacher Competency Tests as Educational Genocide for Blacks: The Florida Certification Examination, Negro Educational Review, 35, 70-77.

Hyman, R. (1984). Testing for Teacher Competence: The Logic The Law, and the Implications, Journal of Teacher Education, 35, 14-18.

In Brief, (1983). The Chronicle of Higher Education, p. 2.

Ivie, S. D. (1982). Why Black Students Score Poorly on the NTE. High School Journal, 65, 169-175.

Kachak, D. (1984). Testing Teachers in Louisiana: A Closer Look, Phi Delta Kappan, 65, 626-28.

Kelley, M. F. (1982). Report of: Arizona's Performance-Based Teacher Certification Program, Arizona Department of Education.

Kline, M. (1973). Why Johnny Can't Read, New York: St. Martin's Press.

Lanier, J. E. (1981). A Futuristic View for Assessing Competence in Teaching, in S. G. Boardman and M. J. Butler (Eds.) Competency Assessment in Teacher Education: Making it Work, Washington, D.C.: American Association of Colleges of Teacher Education.

Legislative Review. (1980). 10, (2).

McDaniel, T. R. (1977). The NTE and Teacher Certification. Phi Delta Kappan, 59, 186-188.

MacDonald, F. J. (1980). Testing Teachers to Ensure Competency: The State of the Art, A paper presented at the annual American Educational Research Association Convention, Boston, Massachusetts.

Medley, D. M. (1978). Alternative Assessment Strategies. Journal of Teacher Education, 19, 38-42.

Medley, D. M. (1981). The Role of Teacher Educators in Teacher Competency Assessment, Inservice: National Council of States on Inservice Education, Syracuse University, 9-12.

Medley, D. M. (1982). Teacher Competence Testing and the Teacher Educator, Charlottesville, VA: Association of Teacher Educators and the Bureau of Educational Research, School of Education, University of Virginia.

Mendoza, A. and Kossack, S. (1982). The State of Florida: A Case Study in Governmental Assessment Legislation, Action in Teacher Education.

Mercer, W. (1983). The Gathering Storm: Teacher Testing and Black Teachers, Educational Leadership, 41, 70-71.

Merrit, R. (1980). The Relationship of Performance in Professional Education Courses to NTE Scores, Educational and Psychological Measurement, 40, 1207-1209.

Montgomery, J. (1979). Can 'Teach' Teach? New York Times.

Nassif, P. M. (1978). Standard-Setting For Criterion-Referenced Teacher Licensing Tests, A paper presented at the annual meeting of the National Council on Measurement in Education: Toronto, Canada.

Nassif, P. M. and Elliot, S. M. (1982). Teacher Certification Testing Technical Challenges: parts I and II, paper presented at the annual meeting of the National Council on Measurement in Education: New York, New York.

Nassif, P. M. (1979). Setting Standards, In Final Program Development Resource Document: A Study of Minimum Competency Testing Programs, National Institute for Education, Washington, D.C.

Nassif, P. M., Pinsky, P. D., and Rubinstein, S. A. (1979). Generating Parallel Test Forms for Minimum Competency Exams, Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, California.

Nassif, P. M., Pinsky, P. D., and Rubinstein, S. A. (1980). Further Work Developing Parallel Tests by P-Value Substitution, Paper presented at the annual meeting of the National Council on Measurement in Education, Boston, Massachusetts.

1983-84 Bulletin of Information. NTE Programs. (1983-84). Princeton: Educational Testing Service.

National Teachers Exam Policy Council. (1979). Guidelines for Using NTEs. Princeton: Educational Testing Service.

National Education Association. (1982). A Closer Look at Teacher Competency Testing. NEA Reporter, January/February.

National Assessment of Educational Progress. (1981). Beginning Teachers Must Make the Grade. NAP Newsletter.

National Assessment of Educational Progress. (1982). Be Wary of Politics of Testing, Educators Told, NAP Newsletter.

National Assessment of Educational Progress. (1982). Testing Teachers: Current Approaches, NAP Newsletter.

Novick, M. (1981). Federal Guidelines and Professional Standards, American Psychologist, 36, 1036-47.

Palaich, R., Flakus-Mosqueda, P., and Wilkins, J. A. (1985). Survey of State Teacher Policies. Denver: Education Commission of the States.

Piper, M. K. and O'Sullivan, P.S. (1981). The NTE: Can it Predict Classroom Performance? Phi Delta Kappan, 63, 401.

Pittman, J. (1975). Actions Taken by State Departments of Education in Developing CBTE Certification Systems. Paper presented at the Association of Teacher Educators' annual conference, New Orleans.

Pugach, M. and Raths, J. (1983). Testing Teachers: Analysis and Recommendations, Journal of Teacher Education, 34, 37-43.

Priestley, M. (1982). Variations in Approaches to Assessment. A paper presented at the annual meeting of the National Council of Measurement in Education: New York, New York.

Pugach, M. and Raths, J. (1982). Testing Teachers: Analysis and Recommendations, Journal of Teacher Education, 33, 13-17, 19-21.

Quirk, T. J., Witten, B. J., and Weinberg, S. F. (1973). Review of Studies of Concurrent and Predictive Validity of the NTE. Review of Educational Research, 43, 89-113.

Sandefur, J. T. (1983). Teacher Competency Testing: The Public's Mandate, unpublished manuscript, W. Kentucky University, College of Education, Bowling Green.

Sandefur, J. T. (1983). Competency Assessment in Teacher Education. In D. W. Jones (Ed.) Governance and Leadership in Teacher Education, Muncie, Indiana: Ball State University. 35, 133-8.

Savage, D. G. (1983). More than Half on Two Campuses Fail Teacher Test. Los Angeles Times (Part 2).

Smith, G. P. (1984). The Critical Issue of Excellence and Equity in Competency Testing, Journal of Teacher Education, 35, 6-9.

Soar, R. (1981). Measures of Quality in the Classroom. Unpublished manuscript, University of Florida, Gainesville.

Summers, A. A. and Wolfe, B. L. (1975). Which School Resources Help Learning? Efficiency and Equity in Philadelphia Public Schools, American Economic Review, 67, 639-652.

U.S. Department of Health, Education, and Welfare, Public Health Services. (1977). Credentialing Health Manpower, (DHEW Publication No. (06) 77-50057). Washington, D.C.: U.S. Government Printing Office.

United States Equal Employment Opportunity Commission. (1966). Guidelines in Employee Selection Procedures. (No. 029 C.F.R.: 1607) Washington, D.C.: U.S. Government Printing Office.

Vlaanderen, R. B. (1984). Teacher Competency Testing: Status Report, Educational Measurement: Issues and Practice, 1, 17-20; 27. (ERIC Document Reproduction Service No. ED 231 814)

Vlaanderen, R. (1979). Trends in Competency Based Teacher Certification, Unpublished manuscript. Education Commission of the States: Denver, Colorado.

Vlaanderen, R. B. (1981). Testing for Teacher Certification: Summary Chart. Unpublished manuscript. Educational Commission of the States: Denver, Colorado.

Vorwerk, K. E. and Gorth, P. W. (1982). Common Themes in Teacher Certification Testing Program Development and Implementation. Amherst, MA: National Evaluation Systems, Inc.

Weaver, W. T. (1978). Educators in Supply and Demand: Effects on Quality. School Review, 86, 552-593.

Weaver, W. T. (1979). In Search of Quality: The Need for Talent in Teaching. Phi Delta Kappan, 61, 120-32.

Weiler, H. (1982). Education Public Confidence, and the Legitimacy of the Public State: Do We Have a Crisis? Phi Delta Kappan, 64, 9-14.

Weiler, H. (1982). Failing Teachers. The Miami Herald.

Weiler, H. (1980). Help! Teacher Can't Teach. (1980). Time. 54-63.

PRINCIPALS AS INSTRUCTIONAL LEADERS: NOW IS THE TIME!

JAMES L. LAFFEY
James Madison University

"It was the best of times, it was the worst of times, it was the age of wisdom, it was the age of foolishness, it was the epoch of belief, it was the epoch of incredulity, it was the season of Light, it was the season of Darkness, it was the spring of hope, it was the winter of despair, we had everything before us, we had nothing before us, we were all going direct to Heaven, we were all going direct the other way - in short, the period was so far like the present period, that some of its noisiest authorities insisted on its being received, for good or evil, in superlative degree of comparison only." (Dickens 1925)

This quote from "A Tale of Two Cities" appears to be reflective of the state of knowledge concerning the principal's role in school reading programs. It is the best of times because we know much more about principals' knowledge of reading and how they can influence reading achievement and reading program development. It is the worst of times because there is no consistent and direct action being taken to implement some of the applicable research findings about principals for the purpose of reading improvement.

Principals - Preparation and Practices

Recently, the Research Committee of the Virginia State Reading Association, which this writer chaired, studied the preparation and practices of principals in reading and reading program development in the schools in the state of Virginia. The results of these studies, which were conducted over a five year period, are most revealing. The reader needs to be cautioned, however, not to generalize results of these studies to principals in schools across the nation.

It is readily recognized by most educators and administrators that leadership in reading requires specialized technical knowledge about reading instruction, reading materials, and reading program development. However, the data collected in these studies, revealed that Virginia's principals are ill-prepared for the role. Among the elementary principals, 21% had not had a reading course before becoming principals. In other words, one elementary principal out of every five never had a reading course. These figures are particularly disturbing since reading instruction in the elementary school forms the foundation for students' success for the remainder of their school years. Equally disturbing was the data revealing that more than two thirds of the middle school and high school principals had never had a reading course before becoming principals. Since principals are responsible for instructional leadership in a school, and since they lack preparation for leadership in the area of reading, it is little wonder that the illiteracy rate in Virginia is so high (over 30% according to recent reports).

While the above figures are alarming, the statistics relative to reading courses that principals have taken since they have become principals are even more alarming! Forty percent of the elementary principals, over 66% of the middle school principals, and 72% of the high school principals indicated they had not taken a reading course since becoming principals. Of course, it is not entirely the fault of the principals for this lack of preparation. Part of the blame can be attributed both to the institutions that prepare principals and the certification boards that bestow credentials on principals. Nonetheless, the principals are also deserving of some of the blame for their lack of preparation, especially in view of the reality that in the surveys it was revealed that the principals, since assuming their positions, made very little effort to acquire the necessary knowledge for their leadership role in reading.

Without reviewing all the details of these studies, it seems pertinent to point out that the principals in the state of Virginia give lip service to the "idea" of leadership in reading program development, but when it comes down to actual practice, their leadership in reading program practices are wanting to say the least. For example, most of the principals (over 90% across school levels) in the Virginia studies stated that reading was the most important curricular area in the school program, yet, vast numbers of them had not taken a single reading course since becoming principals (elementary - 40%, middle school - over 66%, high school - 72%). When these figures are considered in light of other statistics - i.e. that better than 64% of the middle school and high school principals have never taken any reading course - then these facts are most distressing! If American business and industrial leaders were as ill-prepared to assume leadership roles in their positions - and refused to take any action to rectify the sad state of preparation as appears to be the case with many school principals and administrators - then other leaders would remove them from their positions. However, in Virginia and elsewhere, it is apparent that this will not happen. It appears ludicrous to have to admit that in American Education we are permitting the inept and ill-prepared to remain in leadership positions with control of not only millions of dollars of our educational resources, but also the present and future educational lives of our children, yet at the same time we are demanding that teachers be better prepared as they enter classrooms (Kossack 1985) and continue their education for longer periods of time after they enter classrooms. Should this sad state of affairs exist? Quite simply, it should not exist because there is ample evidence in the educational literature that principals and administrators who desire to take an active role in reading program development can make a difference relative to student reading achievement levels and teacher attitudes toward reading program leadership practices.

Reading Achievement and Effective Leadership Behaviors

In order to review pertinent research which has a high degree of credibility in terms of identifying leadership behaviors relating to increased reading program effectiveness, criteria were established to select research which appeared to have some internal and external validity. The four criteria established involved the following types of studies: 1) those which used valid statistical procedures to identify schools with high reading achievement test scores, 2) those in which school effectiveness was defined in terms of student reading achievement test scores, 3) those which identified principals' leadership behaviors that made a difference in students' reading achievement test scores, and finally, 4) those which employed an in-depth case study approach to analyze the relationship between the principals' leadership behaviors and student reading achievement test scores. In each type of study, it should be noted that the most significant aspect was the increased efficiency and effectiveness of the schools' reading programs as indicated by an operational definition of reading achievement that clearly indicated higher and often superior reading achievement test scores as one of the indicators of a successful reading program.

In a summary of the reading research literature reported in the 1970's, (Laffey 1980) analyzed a series of studies which noted the principals' general impact on student achievement and program development. In one study he noted that principals' and teachers' attitudes significantly affected mean achievement test scores of Title I project participants. In a second study which investigated the effects of principals' commitment to improving the reading program as it related to students' progress in reading, it was found the higher the score of the principal on a commitment questionnaire, the higher the students' scores were on standardized reading - achievement tests. In a third study reported in the same summary, it was noted that principals who were recognized by teachers as educated instructional leaders in reading facilitated noticeable improvement in reading achievement for pupils in their schools.

These reports were the forerunners of some more recent research on the administrators'/principals' role in the reading program. This research attempted to identify sepcific kinds of administrative behavior which contributed to higher student achievement in reading and also attempted to relate specific kinds of principals' behaviors to reading program success.

The specific kinds of administrative behavior which contributed to student and school success in reading according to some of the reports are as follows: Creating a healthy, constructive, safe, and achievement oriented school environment was identified in several studies as a necessary ingredient for successful reading programs. Strong administrative leadership was highly correlated with student academic performance (Weber 1971, and Edmonds 1979). Essential to this strong leadership ingredient is the principals' active participation in the on-going school program through 1) observations and participation in classroom instruction (Nash 1983), 2) careful monitoring of student progress (Nelson 1983), 3) having high expectations for student success, and 4) establishing clearly stated goals and objectives for the reading program.

Effective vs. Non Effective Principals

Over the past 15 to 20 years, there has been a concerted effort on the part of reading educators, researchers and educational writers to upgrade the leadership performance of principals in the public schools. Sanders (1971) attempted to define the roles of the administrator by noting 1) that he must be the impetus for the reading program by emphasizing that a philosophy of reading be defined, 2) that he be a participant in learning about reading (in-service) and the decisions made about the reading program, and 3) that he be a facilitator for implementing any new ideas for the reading program. What is most critical, however, but less apparent in this statement is that the principal by taking such an active role is defining his/her leadership role in the reading program. And this, according to a number of different writers, is probably one of the first and possibly one of the most critical tasks to be undertaken if the principal is going to assume real leadership in the reading program.

Nelson (1983) conducted an investigation of eight in-depth case studies of principals' leadership behaviors and found that reading program effectiveness was enhanced by principals who: 1) provide assertive leadership by emphasizing instruction, 2) establish high student expectations, 3) measure and monitor student progress, 4) promote a healthy, well-organized, constructive school climate, 5) establish school goals and objectives and 6) provide for a coordinated instructional program.

Venezky (1979), in an intensive case study of curricular leadership in reading, compared the effect of leadership styles of two urban principals on school reading program achievement. He found that a primary trait of a successful principal was an achievement orientation. "This means that the principal must have a strong desire to have high reading scores and must transmit this desire to staff, parents, and pupils." (p. 19)

To some extent this emphasis on achievement orientation is further emphasized by the successful principal in ensuring that the reading program is well-coordinated. This is reflected in the school's reading program when the goals and roles are well-defined, communicated, and accepted by the school's personnel. Finally, he noted that instructional efficiency which consisted of adaptability of instruction (i.e. the degree to which each student is receiving instruction that he or she needs) and consistency of instruction in which coordination of reading instruction within a building is focused on the student needs across the curriculum and across school personnel so that regular reading instruction is reinforced throughout the school and curriculum. In concert with instructional efficiency, of course, he noted that in successful schools, there was careful monitoring of student progress. The principal received student progress reports periodically and raised questions when progress appeared to be slower than expected.

Kean et. al. (1979) conducted a study to determine what makes a difference in teaching children to read. They gathered extensive amounts of information about the principal, the reading teachers, classroom teachers, reading aides, and fourth grade students in twenty-five schools in and around Philadelphia, Pennsylvania. They then used multiple regression analysis to determine which of 245 variables relating to the five major categories of variables made a significant difference in terms of student achievement. Two variables made a difference. Where principals had experience in the field of reading, the pupils achieved better in reading; and the more the principals observed in the classroom, the better the pupils achieved in reading.

McQuire (1981) conducted a study of teacher perceptions of the principal's actual and ideal role in the reading program in ten schools in Colorado. He found in schools where reading achievement was high, the principals 1) expressed a positive knowledge of the school's reading program, 2) provided fiscal support based on staff needs and 3) conducted school wide evaluation of student progress. Furthermore, he found that in schools where reading achievement was higher, 4) intensive diagnosis and periodic assessment of student reading strengths and weaknesses was conducted, 5) teachers were aware of the school-wide philosophy for teaching reading, 6) a cooperative working arrangement was promoted among the professional and para-professional members of the staff and 8) among the teachers a positive attitude was noted toward the principal's active role and support of the reading program.

Several other studies both reinforce some of the aforementioned leadership ideas and also extend and elaborate on them. A number of researchers reinforced the idea that an effective principal has the ability to provide leadership for the development and implementation of clear and precise goals and objectives by insuring that there is consistency between school goals, grade level and classroom objectives, instructional content and process, and evaluation of student performance (Wellisch 1978, and Manasse 1982).

Other studies cite the importance of the principals' recognizing that instructional improvement is an on-going process (Wellisch 1978, Little 1981) in which they are continually involved by communicating their expectations for instructional improvement to their teachers and by rewarding teachers for their efforts.

Evaluating the Administrator Role in School Reading Programs

Research on effective schools and school achievement indicated that the school administrator/principal is critical in determining both student achievement and school success. Since there is a wealth of evidence which clearly indicates the importance of the principal's role, it seems justifiable that school administrators be evaluated in light of what is known about their effectiveness. Furthermore, it seems that any evaluation should focus on the specific aspects of the administrators'/principals' behavior which has been found to be directly related to student success and/or a school's reading program success. In addition, with due consideration for the research on leadership qualities of principals, it seems important to note some of the research that has been reported on the evaluation of principals. Sanacore (1977) reported on a project in a Long Island school district in which teachers and a reading consultant, with the support of the school district superintendent, devised an evaluation strategy and instrument to be used by teachers to evaluate the school building administrator. The instrument does not evaluate all aspects of the reading program but focuses mainly on important aspects of leadership. Fech and Micetich (1977) conducted a series of workshops on reading program development for principals. The final phase of their project was a request by project directors to have the principals evaluate their performance relating to each portion of the inservice. This evaluation focused on the inservice programs rather than on the impact of the programs on the principals' behavior relating to reading development. McQuire (1981), using a 40 item Likert scale, had teachers and principals assess the principals' role in influencing reading achievement by comparing their actual role in the reading program with the principals' perceived ideal role (what should be). The 40 statements on the scale were directly related to the reading research literature on reading program success and student reading achievement. The results of this study of teachers and principals ranking the most significant principal behaviors relative to reading achievement and reading program development is probably one of the most useful in terms of identifying significant items for inclusion in any evaluation of principals' leadership behavior in reading. Several of the items included in the list of evaluation questions are items which deal with (1) the principal's knowledge of the philosophy of the reading program (2) the principal's strong interest in the program (3) a consistent and continuous instructional reading program across grade levels directed by the principal (4) an active, and supportive communication system relating to the reading program established by the principal for teachers, parents and all school staff members (5) items relating to the principal's allocation of adequate funds to meet the students needs at all grade levels (6) items which concern involving all the teachers in evaluating the reading program and finally (7) items which deal with the principal maintaining a high level of teacher morale. These items as well as many others included in this evaluation project emanate directly from the research literature on reading program leadership.

The Stone Drop Analogy and Ripple Effect

The following illustration of a stone dropped in a pool of water is reflective of the principal's influence in the school reading program. The principal, like a stone dropped in the center of the pool, transmits energy from the central point of impact (his office) in an outward direction, thus influencing all that comes under his administrative influence. And if specific behaviors can be defined as in Figure 1 (i.e. developing a philosophy of reading, defining his role in the program, etc.) then there is much more likelihood that an effective reading program will be developed in the schools.

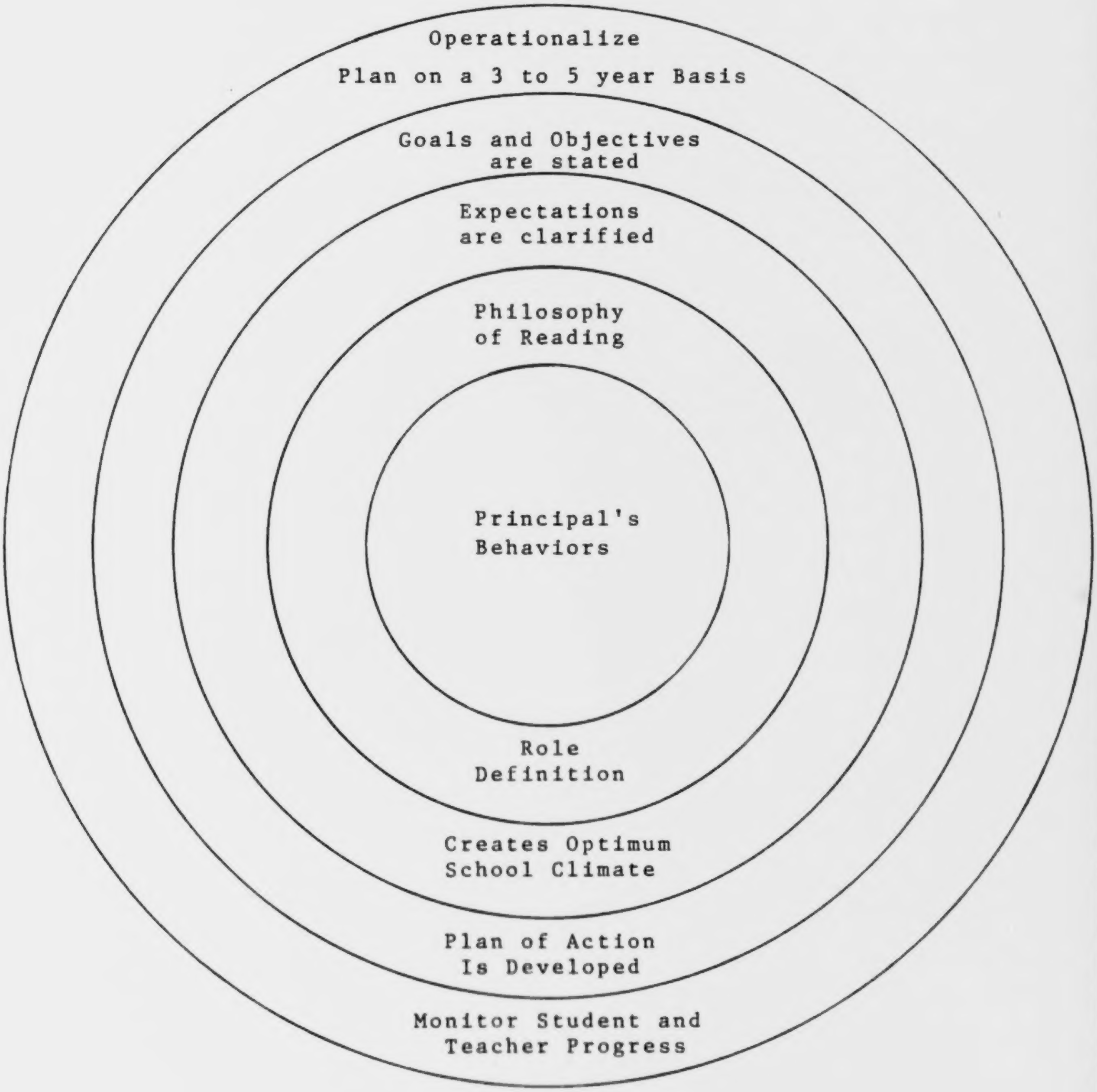

Operationalize
Plan on a 3 to 5 year Basis
Goals and Objectives are stated
Expectations are clarified
Philosophy of Reading
Principal's Behaviors
Role Definition
Creates Optimum School Climate
Plan of Action Is Developed
Monitor Student and Teacher Progress

Figure 1. Specific Behaviors of Principals in Developing Effective Reading Programs

However, it appears that before reading programs can be improved as a direct result of school principals' actions, specific types of research and evaluation projects should be undertaken. First of all, studies which focus on the impact of evaluating principals in terms of specific leadership behaviors relating to reading instruction should be conducted. The focus for such studies should include evaluations from at least two perspectives. Teachers should be given the opportunity to evaluate principals from their point of view regarding their leadership behaviors. Secondly, studies should be conducted from the supervisor and/or central office staffs' perspective of the principals' leadership behaviors.

Some other studies should deal more clearly with defining the leadership behavior of principals. For example, instructional leader is a key phrase that is used in the literature on reading achievement and school effectiveness. Yet, it is not clear exactly what the principal does or should do as an instructional leader in the reading program.

Finally, expectations both in terms of what the principal expects of teachers and students need to be investigated and refined. Even though in a strict sense expectations are not leadership behaviors per se it is clear that they do play a role in school reading program success. But it is not clear as to the specific role the principal plays on either accelerating or depressing both teacher and student responses to the principal's expectations.

Regardless of the focus, there must be new efforts in research, evaluation, and practice which involve the principal as instructional leader. Otherwise the tremendous advances in knowledge of reading instruction which have been made during the last two decades will serve no purpose because of our failure to prepare principals for their critical leadership role in the school's reading programs.

REFERENCES

Dickens, C. (1925). A Tale of Two Cities. New York: Dodd Mead and Company, p. 3.

Edmonds, R. (1979). "Effective Schools for the Urban Poor." Educational Leadership 37, 15-27.

Fech, E. B. and Micetich, B. J. (1977). "The Principal's Role in the Reading Program." ED 153195.

Kean, M. H., Summers, A. A., Raivetz, M. J. and Farber, I. J. (1979). "What Works in Reading?" Summary of Joint School District/Federal Reserve Bank Empirical Study in Philadelphia. Office of Research and Evaluation, pp. 1-67.

Kossack, S. (1985). "Nationwide Use of the N.T.E. - A Survey." A presentation Made at the American Reading Forum Conference, Sarasota, Florida.

Laffey, J. L. (1980). "The Role of the Elementary School Principal in the School Reading Program: Research and Practice." The Reading Teacher 33, 632-634.

Laffey, J. L. and Kelly, D. G. (1983). "Survey of Elementary Principals." Reading in Virginia, Special Edition, James Madison University, Harrisonburg, VA., pp. 3-30.

Laffey, J. L. and Laffey, D. G. (1985). "Middle School Principals Survey." Reading in Virginia, James Madison University, Harrisonburg, VA., pp. 1-32.

Little, J. W. (1981). School Success and Staff Developments: The Role of Staff Development in Urban Desegregated Schools. Final Report. National Institute of Education.

Manasee, A. L. (1982). "Effective Principals: Effective at What?" Principal 61.

McQuire, M. R. (1983). "The Role of the Elementary School Principal in Influencing Reading Achievement." Resources in Education ED 227469.

Nash, M. A. (1981). "A Secondary School Administrator's Guide to Evaluating Reading Instruction." NASSP Bulletin Washington, D.C., pp. 51-54.

Nelson, D. (1983). "School Leadership and Reading Achievement: Eight Case Studies." Paper Presented at the IRA Conference in Anaheim, CA. Resources in Education, ED 273961.

Sanders, P. (1971). "Impetus, Participant, Facilitator. A Definition of the Administrator Role." Journal of Reading 20, pp. 547-552.

Sanacore, J. (1977). "Evaluating Administrative Performances in the School Reading Program." Journal of Reading 20, 312-316.

Venezsky, R. and Winfield, L. F. (1979). "Schools That Succeed Beyond Expectations." Technical Report No. 1, University of Delaware Studies on Education. Resources in Education. ED 177484.

DIAGNOSIS OF READING PROBLEMS: WHAT ARE THE ISSUES? ISSUES RELATED TO TESTING

WAYNE OTTO

University of Wisconsin-Madison

The proposal for this Problems Court Session summarized my position: Otto believes that most formal testing is a waste of time, that most testing behavior is ritualistic and pointless, and that testing should be informal and continuous. Weeks later, when I sat down to prepare my statement for the session, I soon realized that (a) fortunately, my position hadn't changed, and (b) unfortunately, I'd pretty well said it all when I summarized for the proposal. Then the blank yellow pad in front of me got bigger and blanker when I realized that not a single person in the entire American Reading Forum was likely to disagree with any of that. In my mind I heard hoots of ridicule and scorn and I saw my issues slipping away as I tried to elucidate the obvious.

Traumatizing as it was, my reverie helped me back to reality. If everybody can agree--at least in the relative privacy of a professional forum--that most formal testing is pointless and that continuous informal testing is desirable, then why does formal testing and so much of the silliness that goes with it continue? Why do we continue to talk about "reading at grade level"... or above... or below it? Why do we give so-called diagnostic tests that yield no usable diagnostic information? How come we almost never express concern over whether so-called achievement tests test what we teach--or even concern for teaching what's to be tested? Why do we continue to use--or to teach our students to use--tests that guys like Ysseldyke and Algozzine (1982) continue to show to be technically inadequate?

The real issues related to testing, it appears, have much less to do with acknowledging defects and limitations and much more to do with facing the forces and folkways that perpetuate testing practices known to be unproductive or even counterproductive. Such a suggestion is very much in line with the gist of discussions I've had with teachers. By and large those teachers have no delusions about the validity or the usefulness of the tests they give or the test scores they see. When I ask them how they use test scores, their answers tend to be vague--"talking to parents," "reporting to the school board" (or the principal or the program director or some such), or, less frequently, "for grouping," or, more frequently "for making a referral." Almost never do they say they use or even suggest they could use the scores to make instructional decisions. That tends to be so whether they are talking about achievement tests or so-called diagnostic tests.

The teachers I talk with feel (a) guilty because they can't figure out how to use the test scores that are available to them, and (b) resentful because giving the tests consumes time and resources that could be used productively. Yet tests continue to be given.

At issue, it seems to me, is the practice of giving formal tests and, more fundamentally, the array of expectations related to testing. In other words, two basic questions are: (1) Is there any justification for giving formal tests? and (2) Is there a way to establish more realistic expectations regarding the use of test scores?

At issue, too, is the matter of continuous, informal testing. If most of us could agree that continuous, informal testing is potentially productive and desirable for diagnostic purposes, then why is there so little of it in actual practice?

Personally, I think it's because there is a powerful secret society dedicated to perpetuating silliness in testing and to preventing sensible approaches to diagnosis, but my psychiatrist told me never to talk about that. So I invite your discussion.

Reference

Ysseldyke, J. E. and Algozzine, B. (1982). Critical Issues in Special and Remedial Education. Boston: Houghton Mifflin.

DIAGNOSIS OF READING PROBLEMS: WHAT ARE THE ISSUES?

ISSUES RELATED TO LABELS

KAY CAMPERELL
Utah State University

Otto suggests that the real issues concerning formal testing are the folkways and forces that perpetuate use of the tests. I will focus in this paper on some of these forces and folkways. Specifically, the purpose of the paper is to suggest that common practices associated with the use of test scores have serious consequences on the lives of students. Perhaps, as Otto claims, teachers may not use test scores to make instructional decisions. Nevertheless, the educational system does. The scores are used to label students according to the cognitive characteristics supposedly measured by the tests, and, thereby, to assign them to particular programs of study (Leinhardt, Bickel, & Palley, 1982). Moreover, government officials then use the scores and classification schemes to allocate money for the special treatments the students are supposed to receive. These programs and the labels associated with them, not the test scores, may influence the cognitive skills and knowledge students acquire as well as the type of instruction they receive.

This paper is divided into three sections. In the first section I review briefly some of the forces and folkways associated with formal testing in American schools. Next, I review funding policies that perpetuate use of test scores and labels to classify students into special groups. In the last section, the focus is on how labels enable teachers, administrators, and even parents to evade the social and emotional effects the labels have on students.

Testing Folkways

Standardized tests became popular at the turn of this century as a tool for evaluating the success of schools, school programs, and school systems (Resnick & Resnick, 1977, 1982, 1985). Social efficiency policies dominated school administrative practices, and formal tests appeared to provide a "scientific" basis for determining the effectiveness of schools and for assigning students to classes and programs of study. Many educators and the public bought into the illusion of scientific precision, and the notion that test scores could be efficient (i.e., psychometrically reliable) without being valid was not an issue to most people. After all, how could numbers resulting from a scientific test be inexact? An IQ score of 89 sounded precise, as did a reading level of 5.7. Moreover, at the time, it was considered only humane to identify students who might have difficulty learning in regular classes so that they could be placed in classes more suited to their intellectual capacities (Cohen & Neufeld, 1981; Matarazzo, 1972).

Some reading educators, like Gray in the 1920's and Betts in the 1930's did criticize use of standardized tests to assess reading performance (Liebert, 1971). They argued that reading ability could not be reduced to a single score and that diagnosis without the opportunity to observe a student's reading performance was invalid and that such procedures could not convey the complexity of reading levels students achieve or the sources of a student's difficulty in learning to read.

Thus, many reading educators turned to informal inventories as a means of assessing reading performance. Unfortunately, the levels established by such tests became converted into scores also, and commercial rather than teacher-made instruments came to dominate the field (cf. Jongsma & Jongsma, 1981; Powell, 1971). To compound the problem, publishers of standardized tests began to produce tests they claimed could provide diagnostic information to teachers, and use of these tests is even more pervasive in the field than that of informal inventories. Somehow reading educators became seduced by the appeal of scientific precision and administrative efficiency.

Various criticisms of testing practices in reading and standardized testing in general have been published throughout the years. Unfortunately, many of these criticisms, especially those aimed at standardized tests, are off target. It is the users of test results not the tests that are at the center of the problem. Standardized tests were developed to provide administrators and the public with some way to gauge the success (efficiency) of schools, evaluate programs, and compare different school systems. That is, standardized tests originally were developed to monitor the performance of institutions, rather than assess or guide the learning of individuals (Resnick & Resnick, 1985).

Moreover, standardized tests, even those purporting to be diagnostic, are deliberately constructed so that they are not tied to a specific curriculum. This is done both to afford comparisons among schools and to ensure the commercial success of the test. Thus, the results of standardized and commercial tests cannot and should not be used by teachers to shape an instructional response. If teachers feel compelled to do so, it is either because of how they were trained or because of expectations placed on them by school administrators.

Persistent use of test scores for establishing an instructional program reflects, in my opinion, a facet of the methodological superstructures or folkways that have evolved in reading education throughout the years. We perpetuate the practice by training teachers to use the scores to categorize students into groups for instructional purposes and by expecting them to learn labels for talking about students with reading difficulties. We then train them in a diagnostic-prescriptive model which further compounds the problem. Teachers learn from us to seek simple causes of reading difficulty (e.g., skill or knowledge deficits) and to employ simple solutions (e.g., direct instruction, drill and practice).

Forces That Impel Use of Tests

Testing and labels did not arise in a social vacuum. They are tools that serve the needs of various groups in society. Three main groups served by tests and labels are professional educators, funding agencies, and publishers. Numerous labels and formal testing instruments now mark professional boundaries in education. Funding agencies use labels to allocate financial resources to schools, and publishers use them to create and advertise a commercial product. The issue that categorical labels stigmatize individual children and adolescents is far removed from the concerns of these groups.

Formal tests emerged at a time when schools faced a genuine need to cope with increased enrollments and student diversity. Some system for identifying students who required special services needed to be devised. Labels, such as "remedial reader," "learning disabled (LD)," or "educable mentally retarded (EMR)," grew out of sincere concerns about students who did not do well in school. Unfortunately, the various interest groups that evolved due to this concern have magnified testing and labelling practices.

Of particular concern is the impetus to increase use of tests and labels due to federal demands for providing and evaluating special educational services for students who need them. The labels, remedial, LD, or EMR, are categories used by both the federal and state legislatures to allocate funds to school districts (Leinhardt, et. al., 1982; Resnick & Resnick, 1982). These categories are defined on the basis of test scores, and funds are distributed in terms of the number of students who achieve scores that place them within the various categories. Increased financial aid is a powerful force behind schools continued use of tests and categorical labels. The fact that nobody really knows what these tests measure is not at issue. Money is.

In this sense, tests serve a functional need in education, and labels such as disadvantaged, LD, or EMR have an emotional appeal that can be used politically. Nevertheless, the labels are used to ascribe certain characteristics to real students, and it is up to individual educators to employ the labels responsibly. If use of labels negatively influences how students perceive themselves or how others, especially teachers, perceive them, then we need to change how we are conceptualizing individual differences in learning and how to deal with these differences.

Persistence and Evasion

I feel testing and labelling students persists in schools for several reasons. The first concerns expediency. It is easier to administer a test someone else has constructed and to follow their guidelines for interpreting the results than it is to figure such things out for ourselves. Moreover, once a student is classified as a low-achiever or remedial reader, it saves face. Working with these students bolsters the self-image of some people. They see themselves as protectors and defenders of the weak and disadvantaged. Others believe that the students truly are mentally inferior if they can't read, and the labels are just nice ways of saying the students are slow. In either case we evade responsibility to students, as the teacher or the system seems to benefit from the classification schemes not the students.

Another reason these practices persist has to do with money. Billions of dollars are allocated to school systems each year to provide remedial services to students. Schools want and need this money, and I have even seen letters which imply that parents should lie about their incomes so that a system can receive more financial aid. Publishers, of course, profit from tests and instructional materials as do the professors of education who work for them.

Nothing is inherently wrong with schools seeking to increase their financial aid or publishers wanting to make money. The point is, from my perspective, that this aspect of testing exerts a powerful motive for continuing to label students instead of restructuring how they are taught and what they are expected to learn.

Most importantly, however, I think these practices persist because of the education teachers receive in colleges and universities. Most teachers take courses in tests and measurement as well as courses related to learning and cognitive development. Theories related to assessment techniques and learning are somehow rejected by educators, or the theories are taught in such a way that teachers do not understand how to use them. The theory and methodology associated with standardized testing is clear about the proper uses and limitations of these instruments, and it is up to those who teach such courses to make sure that teachers are aware of how to use the theory and technology responsibly.

Closer to home, how we train teachers to think about and provide reading instruction must change. If we continue to teach reading theories, assessment procedures, and instructional methods as practical or putatively "scientific" truths rather than conceptual tools, then nothing will change as teachers will not learn how to apply the ideas we present. It is easier to memorize labels, techniques, and procedures than it is to struggle with the complexity and uncertainty of how to go about teaching individuals to read. We are deluding teachers if we do not expose them to that complexity and uncertainty.

Neisser, in a recent interview, claims that the adjective "intelligence" is one of those words people develop to talk about things they do not understand very well (Goleman, 1983). So, too, are the labels developed to classify students who differ from their classmates in reading. Everyone acts like they know what reading tests measure and what labels for different readers mean; therefore, we avoid dealing with questions about the relationships among the tests and the cognitive capacities they assess as well as what and how students are taught. Once such labels as dyslexic, remedial, or learning disabled are assigned to students, almost everyone relaxes; parents, teachers, administration and the community. Everyone then has an explanation for why learning has not and does not occur. The sign of reading difficulty -- a test score and label -- becomes the signified; the predictor becomes the predicted.

REFERENCES

Cohen, D. & Neufeld, B. (1981). The failure of high schools and the progress of education. Daedalus, 110, 69-89.

Goleman, D. (1983). A conversation with Ulric Neisser. Psychology Today, 17, 54-62.

Jongsma, K. & Jongsma, E. (1981). Test review: Commercial informal reading inventories. The Reading Teacher, 35, 697-705.

Leinhardt, G., Bickel, W. & Palley, A. (1982). Unlabeled but still entitled: Toward more effective remediation. Teachers College Record, 84, 391-422.

Liebert, R. (1971). Diagnostic viewpoints in reading. Newark, DE: International Reading Association.

Matarazzo, J. (1972). Wechsler's measurement and appraisal of adult intelligence (5th ed.). New York: Oxford University Press.

Resnick, D. & Resnick, L. (1977). The nature of literacy: An historical explanation. Harvard Educational Review, 55, 153-177.

Resnick, D. & Resnick, L. (1982). Testing in America: The current challenge. International Review of Applied Psychology, 31, 75-90.

Resnick, D. & Resnick, L. (1985). Standards, curriculum and performance: A historical and comparative perspective. Educational Researcher, 14, 5-20.

DIAGNOSIS OF READING PROBLEMS: WHAT ARE THE ISSUES?
ISSUES RELATED TO PERSONAL CONTEXT

BERNIE HAYES
Utah State University

I have for some time thought that most children learn to read and learn to read quite well. I have also felt for some time that those relatively few children who have difficulty learning to read are not only at a tremendous disadvantage from an educational standpoint but also more disadvantaged because of what happens to them as a result of being "remedial." To rationalize this concern, I need to briefly discuss distinctions between two opposing theories of causes of reading difficulties. I believe that being remedial is perpetuated by education's over reliance on a diagnostic model that could be termed an "ability" model. This model hypothesizes the presence or absence of underlying constitutionally fixed traits such as"perceptual deficits." This model would suggest that poor readers experience difficutly because they process verbal information more slowly than others. This more or less implies a "terminal" condition in the readers. On the other hand, another model could be called a "learning" model. This model suggests that the causes of failure lie more in the instructional environment and makes failure external rather than internal to the learner. This model suggests that a major difference with poor and good readers is in the area of higher mental processes such as strategies and metacognitive behavior.

Considering this second model, Johnston (1985) has, through the use of case studies of adult remedial readers, identified several factors he believes are psychological and social determinants of reading failure. I would suggest that these factors are generally overlooked in mainstream reading theory and contribute greatly to an individual's chronic condition of being a "remedial reader." These factors are combinations of anxiety, attributions, maladaptive strategies, inaccurate or nonexisting concepts about aspects of reading, and goals and motivation.

A brief overview of each of these factors follows for your consideration and discussion.

Conceptual Problems

This factor relates to the substantial conceptual differences that have been found to exist between more and less able readers. If children develop a misconception or fail to learn an appropriate one, any further instruction that presupposes an understanding of that concept will likely be unproductive and perhaps detrimental because of the consequences of the emotional experience related to failure. Clay (1979) has urged that attention be placed on the early detection of misconceptions or poorly developed notions about print and reading.

Strategy

Because the demands for literacy are so prevelant, children cannot ignore reading problems. These children have two alternatives when confronting deficiencies in reading. Attempts can be made to correct the problem, or they can accept it and attempt to cope. Poor readers often strive to devise ways to cope with having a problem in reading and also ways to cope with reading demands. Many of these can be self-defeating. Avoidance of print and an over-reliance on general knowledge and the context to figure out unknown words are examples of self-defeating coping strategies.

Anxiety

The risk of being perceived as stupid by others may cause individuals to systematically avoid reading, thus ensuring that any reading skills they do have will not become automatic through practice. This fear of being perceived as stupid may also produce stress. Such characteristics as resistance to entering the learning situation, fear responses (trembling), inability to avoid making incorrect responses, compulsive responding, regressive behavior, and the loss of the ability to delay a response are a few of these observed in reading-disabled children. Such conditions make learning to read an extremely difficult task.

Attributions

This factor relates to a perception of a cause to which individuals attribute their failure. This perception is associated with attributions of failure to uncontrollable, invariant factors such as lack of ability rather than effort. Reading educators have suggested that many teachers err by holding expectations that are too low, by pacing instruction too slowly, and by ignoring or under-emphasizing the substantive aspects of tasks when teaching "low" groups. In doing so they implicitly tell less able students that they are "constitutionally" less able. This attribution of reading failure to a cause for which there is little hope of remediation leaves little hope for success in learning to read.

Motivation and Goals

Goal conflict, the desire to protect one's self concept exorting a stronger influence than the desire to learn to read, can sometimes be a powerful deterrent to overcoming reading difficulty. A basic difference between motivation in ego-involving and task-involving situations seems to be determined by whether the central goal is to protect or enhance one's self concept, or to complete a task. The results of demonstrating an inability to read is highly ego involving and may cause motivation to shift from the learning task of reading to protecting one's self worth.

Considering the present information regarding how teachers interact with children possessing limitations in reading ability and society's pervasive assumption equating literacy with intellect, it is my position that the above contextual factors undoubtedly play a significant role in contributing to "remedial readers." It appears to me that we may perpetuate this remedial condition for many children by an over-reliance on models of diagnosis that fail to consider the personal context of the individual. Effective diagnosis should involve the examination of the learner's goals, motives and situations.

REFERENCES

Johnston, P.H. (1985). "Understanding Reading Disability: A Case Study Approach", Harvard Educational Review, 55, No. 2.

Clay, M.M. (1979). The early detection of reading difficulties: A diagnostic survey with recovery procedures (2nd ed.). Exeter, NH: Heinemann.

Commentary on PROBLEMS COURT SESSION
Diagnosis of Reading Problems:
What Are the Issues?

Reactor: RUTH J. KURTH
North Texas State University

In describing the history of standardized testing, Resnick (1982) writes:

> Educational testing initially served a gatekeeping function between school grades (a function to which it has recently returned.) Although it pointed to some difficulties with curriculum and instruction, the major function of eductional testing did not turn out to be assisting teachers in their instructional planning. Indeed, there was considerable opposition from teachers to the testing. Support for comparative testing initially came from school administrators, who were particularly pleased about the accountability function of the tests.

The over emphasison the "gatekeeping and accountability function of standardized reading tests" seems to be the focus of much of the criticism of testing discussed in this Problems Court Session, "Diagnosis of Reading Problems: What are the issues?" Otto feels that testing is counter productive when it does not help teachers make instructional decisions. Camperell feels that the use of standardized tests to group children for various placement aids administrators but harms children by labeling them. Hayes feels that assessments which fail to take into account the personal context of each child cannot adequately measure his/her progress.

The writers of the papers presented in this Problems Court all agree that the standardized tests currently in use have certain inherent weaknesses which limit their usefulness for instructional decision-making in the classroom. Johnston (1984) claims that the historical focus which stressed the testing of large groups for comparison and placement fostered the production of a certain type of test. Because of this historical focus, reading tests were developed which covered large amounts of curriculm content, which stressed silent rather than oral reading, which were product rather than process oriented, and which ignored the metacognitive components of reading. In most instances the validity of reading tests were determined from correlations with other reading tests; therefore, new tests differed little from those previously published.

The types of reading tests which have been published are of little value to teachers attempting to make diagnostic decisions in classroom settings. Yet, as Otto notes, teachers feel either resentful or guilty because they can find few uses for data from standardized reading tests. These feelings increase when test writers and evaluators such as Herman and Dorr-Bremme (1983), Ward (1978), and Yeh (1980) claim that teachers know too little about standardized norm reference testing and therefore use these tests improperly in their classrooms. It is relatively easy for test developers to blame improper use of the tests rather than to determine if inherent weaknesses in the tests or in the assumptions guiding their development make them inadequate for classroom decision making.

Johnston (1984) comments on the need for change in reading assessment:

> There is a great need for researchers to reflect upon and communicate the importance of their findings for assessment practice in applied settings. Research findings of late have tended to emphasize the importance of process over product, yet educators and researchers persist in depending on the more conveniently obtained product data. ...many researchers continue to ignore the metacognitive components of reading in their assessment and the importance of metacognitive aspects of assessment for instruction.... Instead of a concern over response outcomes, right or wrong, there needs to be greater concern over the reasons behind the responses. The bottom line is that we need to worry more about the assessment of process in the individual, and about the process of assessment in context. ....Failure to challenge the status quo may be the worst form of intellectual complacency.

The fact that some basic assumptions which have guided the development of standardized reading tests are faulty is an issue that needs to be addressed. Suggestions for reevaluating the development and use of standardized reading tests are slowly surfacing and need to be addressed.

A second issue raised in this Problems Court is that although continuous, informal testing is potentially productive, teachers actually do little of this type of assessment. Continuous, informal diagnosis is clearly dependent upon the observational skill of the teacher. If teachers are to become adept at using continuous, informal assessments, models for this type of assessment will need to be developed and taught. Teachers will need to become sensitive observers and recorders of students' reading behavior. The overreliance on quantifiable test data has caused some teachers to distrust or at least give less credence to the validity of their observations of children in instructional settings. Helping teachers realize the value of careful observational data based in an instructional setting and completed over a length of time would likely increase the amount and quality of informal observations. An effective management system for recording daily or weekly observational data will also need to be developed if continuous assessment is to be done. Teachers will spend time doing activities which they feel most competent to perform. If they trust their abiltiy to do informal, continuous assessment in reading, the amount of informal assessment in reading classrooms will increase.

Continuous, informal assessment by qualified teacher observers appears promising. Data gatherd by qualified teachers in instructional settings should prove to be more useful in making instructional decisions than the data from product based standardized tests now being used. Carefully gathered anecdotal and observational data can shed much new light on the contexts in which children learn. However, this change to a one-to-one descriptive model of reading assessment places the responsibility for the quality of the assessment on the teachers rather than on the test instrument. Teacher training programs will need to teach methods for careful observation of student's reading behaviors in instructional settings. The validity of the assessment will depend, to a large degree, on the observational skills of the teacher-observer; therefore the careful training of classroom teachers will be essential.

The time is ripe for a careful evaluation of the issues of assessment and evaluation in reading. Movement from product to process oriented assessment and from formal to informal assessment can be seen. However, the impetus for improvement in reading assessment will probably not come from current test developers or from current administrators intent on gate keeping and accountability. Leadership in this area must come from reading researchers and educators who are familiar with the development of reading ability in a classroom setting. Discussing these issues in a Problems Court at a Reading Forum is a step in the right direction.

REFERENCES

Gronlund, N. (1985). Measurement and evaluation in Teaching. New York: Macmillan.

Herman, J. and Dorr-Bremme, D. (1985). Uses of testing in the schools: A national profile. In W. E. Hathaway, (Ed.), New directions for testing and measurement. San Francisco: Jossey-Bass.

Johnston, P. H. (1984). Assessment in reading. In P. D. Pearson (Ed.), Handbook of Reading Research. New York: Longman.

Resnick, D. P. (1982). History of educational testing. In A. D. Wigdor & W. R. Garner (Eds.), Ability testing: Uses, consequences, and controversies (Part 2).

Ward, J. Teacher and testing: A survey of knowledge and attitude. ERIC Document No. ED 193-258.

Yeh, J. A reanalysis of test use data. ERIC Document No. ED 205 590.

REACTION PAPER

Commentary on PROBLEMS COURT SESSION
Diagnosis of Reading Problems:
What Are the Issues?

Reactor: CAROL DANA
Marquette University/Wisconsin

Diagnosis Of Reading Disability Through An Interactionist Perspective Model

The panel members of the problems court on the diagnosis of reading disorders raised several issues regarding diagnosis, ranging from the deleterious effects of children being labelled after diagnostic testing to the core question of whether any current testing procedures are at all helpful in the corrective process of reading disorders. Respondents from the floor not only reaffirmed the concerns that were raised by panel members but also initiated discussion on additional problems found in current diagnostic procedures such as basic weaknesses in testing instruments, difficulties in translating test results into procedures for correction, and the lack of communication between the test examiner and the remedial teacher. The consensus of opinion appeared to be a dissatisfaction with current practices being used in the diagnosis of reading disorders and a call for possible alternatives.

My purpose in reacting to the problems court was to draw implications for the practitioner, based on the concerns that were raised during the discussion. Simple solutions to the concerns are not available; a single solution is nonexistent. I, therefore, decided to address one concern that was raised several times during discussion-the limitation in diagnostic reports of the examiner not providing the remedial teacher with clear recommendations for remediation. Countless scores are usually provided from the many tests and subtests that are administered, but information on what the scores mean and on how the scores can be translated into an effective remedial program is not always clear.

In addressing this concern, I feel there must be a change in the current model of reading disability and concur with Lipson and Wixson (1985) who proposed a change away from a search for causative factors and toward the identification of conditions under which the child can learn. They asserted that an interactive perspective of reading disability should be taken with reading being perceived as a dynamic process that is accomplished by the reader in interaction with the text, task, purpose, and setting conditions of the reading situation. They posited the need to identify learning situations in which the learner can and will employ his knowledge and skill to improve his reading proficiency. Such a direction would lead to a diagnostic process that would identify instructional interventions to help remediate the reading weaknesses of an individual student.

To operationalize Lipson and Wixson's interactionist model for reading disability, a new direction in assessment should be taken. Mark Ozer (1977) has developed an interactive assessment model which is consistent with the interactionist perspective proposed by Lipson and Wixson. His procedure has as its focus the identification of a plan for treatment. This plan centers on answering three key questions:

1. What does the child now do successfully; where do you start?

2. What are the goals?

3. What works in accomplishing these goals?

Ozer believes that the process of assessment is a sampling of what works. The examiner is an active participant determining student needs and identifying instructional interventions that are best suited to meet these needs.

In answering question one, the examiner works with the child, his teachers, and parents primarily through interviews to identify areas in reading where strengths and weaknesses can be found. The primary focus of interactive assessment takes place when attempting to answer questions two and three. The examiner, after determining goals for the student, based on the information gathered from question one, takes the role of a teacher and conducts "trial teaching," testing a set of instructional interventions in order to identify the interventions that are best suited to help the student reach the established goals. Throughout this process a dual feedback system is in operation. The student is continually made aware of what the examiner is doing and why she is doing it. And the learner, in turn, is asked to give input on each instructional intervention being used.

In the selection and design of each instructional intervention being tested, the only criteria that Ozer recommends the examiner follow are that the task should be from 10-15 minutes in length and that it can be modified so that success can occur. The intervention design as far as presentation style, materials, and feedback aspects is only limited by the ingenuity of the examiner.

A key element of interactive assessment is the enhanced awareness by the student of how to take responsibility for his learning. The role of the examiner is determining what works for remediation and guiding the student to focus his awareness on the effectiveness of each intervention strategy. The role of the student is to use self monitoring to help identify the interventions that seem to work best for him. During the subsequent remediation program while working with the remedial teacher, the student would gain skill in the use of the interventions and eventually acquire a set of compensating strategies to use independently when approaching a learning problem.

The problems court on issues in diagnosis of reading problems raised many concerns with current practices in diagnosis. In addressing the concern of having the examiner provide the practitioner with a clear plan for remediation, I believe Ozer's interactive assessment procedure offers a diagnostic framework that is consistent with an interactionist perspective and offers the examiner a method for identifying conditions that work in the remediation of reading disorders. The task remains to test the procedure. The reading diagnostician would, first, identify a master set of instructional interventions that could be used to help disabled readers; second, formulate ways in which these strategies could be adapted for use with students having specific learning needs. The third step would be to implement the procedure and work with the remedial teacher to study if interactive assessment can be an effective diagnostic tool to identify interventions that facilitate the remediation of reading disorders.

REFERENCES

Lipson, M.Y., & Wixson, K.K. (1985). Reading disability research an interactionist perspective. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Ozer, M.N. (1977). Assessment of children with learning problems: A child development approach. In W. Otto, C.W. Peters & N. Peters (Eds.), Reading problems: A multidisciplinary perspective (pp. 89-107). Philippines: Addison Wesley Publishing Company, Inc.

A VIEW OF A STATEWIDE EVALUATION: IMPLICATIONS OF TENNESSEE'S CAREER LADDER
UPDATE 1985

JANET R. HANDLER
DEBORAH L. CARLSON
University of Tennessee

This update represents the fourth of a series of short reports issued since Summer, 1984. The series was initiated to provide pertinent, timely information to a broad network of educators and governmental policymakers interested in career ladders and other forms of incentive programs. Each update has focused on a specific phase of Tennessee's Career Ladder Program for teachers, from initial formulation and legislative passage through its first year of implementation. Data presented in these periodic reports has been gathered as part of a 20-month project sponsored by the U.S. Department of Education and conducted at the university of Tennessee, Knoxville.

The Tennessee Career Ladder Program has received considerable attention as one of the first statewide programs to be enacted. It remains one of a select number of such reform efforts to have developed a uniform framework for evaluation and advancement, rather than leaving to local district initiative the creation of the specific incentive program in use. Tennessee's model does include local involvement in evaluation at the Probationary, Apprentice, and Level I stages of the teacher career ladder, subject to state guidelines and approval. Access to the two upper rungs (Levels II and III), corresponding generally to at least 8 years of teaching experience for applicants, is strictly governed by intensive state-conducted evaluation procedures.

During the 1984-85 school year, these procedures moved from the field testing and revision stages into full implementation in the schools. Using a one-time process referred to as fast-tracking, individuals currently teaching and who met basic eligibility requirements could apply for entry to Level I via one of five routes (discussed later in this report). A salary supplement or "bonus" of $1,000 was earned by those attaining Level I status. As of the end of April, new figures for teacher Career Ladder sign up indicated that most teachers eligible for Level I applied (27,072). Fewer of those eligible chose to apply for advancement to the upper two career levels during this first year of implementation (7,743 applied). It should be noted that for future teachers and those whose Tennessee service began with the 1984-85 school year, participation in the Career Ladder Program is mandatory.

Consistent with the objectives of the federally funded case study, several activities designed to portray career ladder implementation were undertaken. Beginning in September, 1984, 18 schools in 6 school systems were selected as field sites for fall and spring interviews involving 150 teachers as well as building administrators. Additionally, a comprehensive survey was mailed to a stratified sample of 2,105 teachers across the state. This instrument, distributed in February, 1985, was used to determine teachers' experiences with the career ladder as well as their reactions to its features and conceptual bases. Results of these two major data collection activities are highlighted in this update, with detailed findings to be released later in the final project report (Part II: Shaping Tennessee's Career Ladder--1985).

Interview Results

The field site interviews were designed to ascertain how the program was being implemented in various types of school districts and how it was perceived by teachers and administrators. Two sets of interviews (Fall and Spring) were conducted in each school, so that key events and changes could be ascertained. Documenting initial reaction and describing developments associated with the early operation of the career ladder program were believed to be especially important for helping those in other states learn from all facets of the Tennessee experience.

The six school districts cooperating in this project were selected to represent a variety of demographic characteristics including geographic location, type of community, and general socioeconomic level. In each district, elementary, middle or junior high, and high school levels were represented, with a total of 18 schools participating. Letters and telephone calls were used to maintain contact between site visits.

A high degree of consistency was evident in the responses obtained across districts. Teachers expressed considerable reluctance to climb beyond Level I of a career ladder that they generally regarded as having an unfinished status. Their collective comments indicated that not being sure what to expect regarding the new statewide evaluation process (e.g., how various components would be assessed) and being uncertain about the responsibilities they would be expected to assume (e.g., under extended contract provisions for Levels II and III) were common sources of concern. These concerns were not substantially different in the Spring interviews when compared with the Fall data.

The pace of implementation and the difficulty of establishing clear statewide communication were underscored through the interview data as perceived areas of concern. Teachers and administrators reported numerous experiences leading to feelings of confusion and frustration. In some cases, the situations perceived as changes in the career ladder program were actually unforeseen special cases requiring decisions about eligibility or comparable issues As situations arose requiring clarification or revision, interview respondents reported receiving delayed or inaccurate information. Their responses indicated that state evaluators and other staff representing the State Department of Education were found to have incongruent information on procedures and requirements. Adding to the reported problem was the fact that local school district officials also became bearers of outdated or inaccurate information often enough to represent a major concern of teachers and principals interviewed.

After these communication problems arose in the Fall, there were energetic State Department efforts to strengthen this critical area. Among those interviewed, communication was still regarded as a key concern in the Spring, particularly in the rural schools visited. While more teachers now mentioned receiving mailings from the State Department and other sources, there were indications that the peer "grapevine" was still a primary information source in these schools.

Interview respondents were generally supportive of the concept of performance evaluation, but their comments about the specific process developed for Tennessee contained several recurrent themes. There was concern that the procedure as carried out by some evaluators did not provide adequate postobservation feedback. There was also a common worry that it would be too easy under the present program for outside evaluators to misjudge a situation or even to be intentionally misled. Negative reaction was most prevalent concerning the portfolio required for Levels II and III, which was to include documentation of specific aspects of performance across a 5-year period. In addition to the preparation time required and the weak relationship to teaching behavior perceived, respondents in nearly all schools reported being concerned that factors beyond their control (e.g., lack of leadership opportunities, inadequate resources, burdensome workload) could unfairly decrease their chances of success.

One of the most positive elements associated with the Career Ladder Program was reported to be the staff development model known as TIM (Tennessee Instructional Model). Although not originally created for the career ladder, its inclusion as a fast-track option for the 1984-85 school year brought many teachers into contact with the TIM training modules. These were found by numerous respondents to provide a useful review or pulling together of management, planning, and other aspects of instruction. The primary criticism voiced in relation to TIM was its adherence to a specific approach that seemed to these individuals to limit its applicability to diverse teaching situations.

When asked during the interviews what critical elements should form the basis for recommendations to other states or districts, the most frequent responses underscored the importance of teacher input at the planning stage. Suggestions focused also on ways to reduce the paperwork and duties that were perceived as burdensome, as well as increasing salaries to levels more appropriate for their training and responsibilities. During the Spring interviews, teachers and administrators reiterated earlier recommendations and stressed in addition the critical need for improved measures to ensure evaluator skill and consistency. Steps taken to clarify purposes, streamline procedures, and slow the pace of implementation were also deemed quite important.

Survey Results

The survey instrument developed for this phase of data collection concentrated on three major areas: communication, the fast-tracking process, and evaluation procedures. Participating schools were selected from each of the 18 Teachers' Study Council regions. This selection process yielded a stratified sampling which resulted in 54 schools equally balanced across grade level categories. Based on faculty counts within the selected school, 2,105 surveys were distributed.

When all returns were received, there were 1,039 teacher responses. The first step asked respondents to assess the quality of communication they had received about the Career Ladder Program. Results indicated a substantial level of concern, with 40% specifying "some concerns" and 23% "strong reservations" [the other two responses options were "generally favorable" (33.7%) and "very positive" (3.3%)]. In order to obtain the most specific and accurate information possible about the communication that took place, those surveyed were asked several questions about common data sources (e.g., State Department Mailings, School System Central Office, Tennessee Education Association Staff). Their responses included: whether they had used each source, whether they obtained the information sought, and whether the information was accurate, as well as examples or comments they wished to provide concerning the specific data sources.

The sources of Career Ladder information that were reported as being used most often were: peers (80.7%), TEA Mailings (76.1%), principal (77.3%), and State Department Mailings (73.6%). School principals were felt to have provided the information sought by the highest percentage (49.1%). Assessments of information accuracy resulted in a similar ranking of data sources. The highest proportions for data sources rated as at least "sometimes" accurate were: peers (62.9%), principal (63.9%), TEA Mailings (62.9%), State Department Mailings (57.6%).

Results concerning communication sources and quality were generally consistent across grade levels and eligible career levels of survey respondents, and closely paralleled total sample results. When written comments were analyzed in addition to the structured items discussed above, a more strongly negative pattern was evident. Of 109 comments made concerning the role of the State Department-Nashville Staff as a data source, for example, 78% pertained to inaccurate, delayed, or in other ways not useful information. Another 15.6% of comments described experiences or reactions related to unavailable information from this data source, which elicited the greatest number of comments concerning communication. The second highest frequency of comments on this subject was 66% in reference to the State Department Mailings as a data source. Here, 90.9% of the comments described communication as not useful, inaccurate, late or unclear, with 7.6% reporting helpful information received. Among the other nine data sources listed, no single category was the subject of over 41 comments. Those for which the positive comments exceeded the negative were: principal, Teachers' Study Council, and TEA Mailings.

A second major focus of the instrument was the fast-tracking process used for 1984-85 to facilitate qualification of eligible teachers for Career Level I. This process, in which 83.3% of the 1,039 survey respondents reported participating, required successfully completing one of five options: National Teachers Examination-Core Battery, National Teachers Examination-Specialty Area Test, Tennessee Career Ladder Test, Staff Development, or Full Evaluation. Regulations governing each option were set forth in terms of cutoff scores, duration and nature of training or assessment experiences, and other necessary aspects. The fast-tracking concept was positively regarded by the survey respondents, with 30.4% indicating their views were "very positive" and 35.5% "generally favorable." Only 7.1% indicated "strong reservations" concerning this concept, with 18% marking "some concerns" and 9% not responding.

Respondents were also asked to indicate whether or not they had experienced each of the five fast-track options, and to list strengths and weaknesses of each. It should be noted that applicants could have participated in more than one option. Staff development was respondents' most commonly chosen option (44.3%), followed by the Career Ladder Test (32.6%), NTE-Specialty Area Test (16.6%), NTE Core Battery (10.2%), and Full Evaluation (8.2%). There was considerable variability across regions of the state in terms of fast-track options selected. Less response diversity occurred, however, when results were analyzed across grade levels or eligible career levels.

In the section where respondents were asked to separately list strengths and weaknesses of each fast-track option, comments were most frequently made regarding the Staff Development option (237). Next in frequency were the Career Ladder Test (203), NTE-Core Battery (161), NTE-Specialty Area Test (150), and Full Evaluation (87). Respondents making comments were most positive about the Staff Development option, with 57.8% stating in various ways that it was appropriate for its purpose, helpful to teachers or a good training experience.

Remarks concerning the other options were more negative in each case. The National Teachers Examination-Core Battery was felt to be an inappropriate measure by 60.3% responding to this item, with another 14.3% critical of its content and 7.5% critical of its length. The Specialty Area Test was similarly regarded, although in this case 25.3% found the content to be a positive feature, 16.7% thought it negative, and 46.0% found the test inappropriate for the purpose. A comparable level of dissatisfaction with the Career Ladder Test was expressed among the 203 respondents who commented to it, of whom 54.7% found it inappropriate for the purpose. When all positive comments concerning this instrument were tallied, they accounted for 11.3% of the total for that option.

The third major portion of the survey dealt with the statewide evaluation process developed for the Career Ladder Program. Before addressing their experiences with each component of that process, respondents were asked to indicate the degree to which they favor the concept of a statewide teacher evaluation process. Overall, results reflected a low level of support for the concept, with 39.2% marking "some concerns," 25.7% "strong reservations," 22.2% "generally favorable," and only 7.0% "very positive" (5.9% no response).

A second general question asked how familiar respondents felt they were with the fit or relationship between local and state evaluation. In the Tennessee program, evaluation for Probationary, Apprentice, and Level I status is conducted by the local districts under state guidelines and monitoring, while the state takes over full for Levels II and III. Most teachers expressed at least general familiarity with these relationships (61.0%), of which 15.5% felt "very familiar."

Respondents' experience with the components of the career ladder evaluation process varied, particularly in view of the fact that the state evaluation teams had not yet completed the full three-visit cycle (including portfolio review) with most Level II and II applicants. The evaluation component with which teachers reported the highest degree of experience was classroom observation (45.1%). Fewer than 10% reported that they had experience with either student questionnaires, interviews, or portfolios in this context.

There were spaces provided on the survey instrument for respondents to list strengths and weaknesses of each career ladder evaluation component. Their comments were analyzed to develop categories and tabulate responses within these. The components eliciting the greatest number of comments were: observation (187), Professional Skills Test (149), peer questionnaire (147), and portfolio (146).

Comments concerning the classroom observation were approximately evenly split between strengths (50.8%) and weak points (49.2%). The strengths most frequently cited dealt with usefulness or appropriateness of the observation and its ability to provide helpful feedback to the teacher. The weaknesses mentioned pertained to its being a poor measure of teaching ability and a technique prone to subjectivity or faulty execution.

Nearly all of the 132 remarks about the student questionnaire used in career ladder evaluation were negative. This instrument was intended as a tool that could focus students' attention on actual behaviors rather than general impressions. Although most respondents had not yet directly experienced this particular procedure, 34.9% felt it was not appropriate, too subjective, or a poor use of time, and another 34.1% judged it to be a weak measure of teaching performance. However, 20.5% listed strengths such as providing useful feedback to the teacher and being a good idea for gathering evaluative data.

Two-thirds of the comments regarding the peer questionnaire, a tool administered to three colleagues (chosen from a list of six submitted by the applicant), identified weaknesses perceived in this procedure. Most of these weaknesses concerned a lack of clear relationship to teaching performance and a high level of subjectivity or potential bias.

Responses pertaining to he principal questionnaire, a form to be filled out concerning the applicant's performance, revealed a distinct split of opinion. While 51.4% of the comments were positive and found this to be an appropriate data source, the other 48.6% of the 103 comments gave opposite views.

The response pattern for the Professional Skills Test as an evaluative tool showed the same type of negative pattern as the student questionnaires. There were fewer than 20% positive comments about this data source, while 45.0% felt it was a poor way to measure teaching ability and another 36.2% indicated other weaknesses related to the test.

The candidate interview was dropped from the career ladder evaluation process shortly before the surveys were distributed. This action resulted from a controversy surrounding the access some teachers had been found to have (largely through TEA workshops) to a field test version of the rating scale used to score the interview. These interviews typically lasted several hours, and teachers frequently reported devoting extensive preparation to assemble documentation for their responses. When survey comments were analyzed, there were nearly twice as many negative (65.5%) as positive (34.5%) responses. Those who were positive felt that the idea was good, and the interview could provide useful information and service a helpful purpose. The largest number of weaknesses identified dealt with the perceived subjectivity, inappropriateness of purpose, and excessive time requirements of the candidate interview as a data source for the career ladder evaluation process.

The final data source was the portfolio of sample lessons, examples of leadership activities, and other specified types of documentation. This technique was the subject of numerous comments, most of which (83.6%) were unfavorable. Respondents felt that the portfolio required too much time (39.0%), was too subjective or contrived in nature to be a useful source of feedback (20.6%), or was an inappropriate measure of teaching ability. It should be noted that at the time of the survey, actual portfolio reviews had just recently begun by state evaluators. Thus, respondents were largely noting strengths and weaknesses based on their experiences in getting portfolios ready to be assessed, or based on their understandings of what the portfolio entailed and why it was being utilized as a data source.

Summary

The results of two major data collection efforts have been highlighted in the preceding sections. Data from (a) 150 interviews and (b) a survey completed by 1,039 respondents are currently being pooled with other important sources of information to develop an overall portrayal of the processes and issues associated with implementing a teacher career ladder in Tennessee. This information, including recommendations and implications, will be presented later in the final project report. It is possible, however, to identify several generalizations based on the two data collection activities reported here:

1. Teachers have tended to trust and rely more heavily on local than distant sources of information about the Career Ladder Program, despite the fact that the program is being directed and shaped at the state level. Early communication problems engendered a negative expectation regarding information received from state level agencies. Although corrective steps have been taken to decentralize and improve communication, these have not been completely successful according to data obtained as late as April, 1985.

2. Fast-tracking, a system built into the Tennessee legislation to permit eligible teachers to gain relatively expedient entry to the career ladder, was favorably regarded as a concept by survey and interview respondents. However, when operational features of the fast-track process were discussed, there was considerably less satisfaction. Given the short time for program implementation and the complexity of the fast-tracking system (e.g., five options with varying procedures, deadlines, time requirements, and costs), communication breakdowns and confusion over rules could be anticipated. These problems proved damaging to overall perceptions about the program.

3. While teachers in the study supported the need for performance evaluation, they seemed quite concerned as a group about having a statewide evaluation process. Their analyses of the Career Ladder evaluation system pointed to several important criteria. Teachers want evaluation tools that show a clear relationship to teaching performance and expectations; are resistant to bias on the part of students, colleagues, or administrators; and are reasonable in the demands on teacher time. Prior to widespread first-hand experience with all parts of the evaluation process, teachers gave mixed assessments of the classroom observation and the principal questionnaire. They identified considerably more weaknesses than strengths in the other evaluation data sources It will be interesting and important for these assessments to be studied further as implementation of the Tennessee Career Ladder Program continues into a second year.

THE ROLE OF THE READING EDUCATOR IN THE TRAINING OF ELEMENTARY SCHOOL PRINCIPALS

RUTH JUSTINE KURTH
North Texas State University

Introduction

It is almost axiomatic in professional education circles that the principal exerts a great influence on the quality of the elementary school reading program. In his discussion of the role of the principal, Georgiades (1985) states: "Leadership at the local building level is a key factor in the improvement of the quality of instruction. A school is but a reflection of its principal." However, reading educators have generally given little attention to the role of the building principal in the reading program. Little research focuses on the role of the principal rather than on the role of the teacher or learner. This lack of interest in the development of competent administrators can be seen in the Program for this year's American Reading Forum "Evaluation in Reading". Although the evaluation of administration formed one strand of a three strand theme; only a small percentage of papers submitted dealt with the evaluation of the administration of reading programs. This Problems Court will address the lack of participation by reading professionals in training elementary school principals who can be instructional leaders in the area of reading. This lack of interest in the development of competent principals who can serve as instructional leaders of reading programs is apparent in three areas: 1. the lack of clearly defined research on the role of the principal in the reading program; 2. the lack of involvement of reading professionals in setting standards for evaluating, certifying and licensing elementary principals; and 3. the lack of college courses designed specifically to help principals become effective instructional leaders in reading.

Problem One

Although research clearly shows the strong influence of the building principal on the quality of the school reading program, research by reading professionals to help us understand the role of the building principal in the reading program is noticeably lacking. The findings of the Ford Foundation in its report, A Foundation Goes to School (1976) are paralleled by the findings of the National Association of Secondary School Principals in the Model Schools Project (1978). One of the key findings of both of these studies was that the most significant person in the change process in schools is the building principal.

In the Beginning Teacher Evaluation Study (1980) conducted by the Educational Testing Service and the Far West Regional Laboratory for Educational Research and Development, "successful" schools were compared to "unsuccessful" schools. A major finding from the research in the "successful" schools is that in productive schools the principal acts as an assertive instructional leader and also assumes responsibility for the evaluation of achievement. In a three year study of school leadership (1976), The National Consortium for Humanizing Education found that the principal has a vitally important role in the school and that, in many ways, he or she models roles and behaviors which teachers and children tend to adopt. A study by the Council for Basic Education (1971) noted that in four inner-city schools that were successfully teaching reading, strong leadership was one characteristic common to all four schools.

However, reading research in the schools has focused primarily on the role of the teacher and very little research has been done concerning the role of the principal in the reading program. It is naive to assume that improving the performance of individual teachers will result in a well balanced program of instruction throughout an entire school. We have little concrete evidence which shows how the behavior of building principals affects the reading program either positively or negatively. Often reading researchers have been content to summarize leadership studies done by others and to generalize these results to the reading program. This summarizing of general studies leads to vague suggestions for administrators which are not really reading specific, such as "a principal must have clear goals for the reading program, a principal must set up a climate conducive for learning, a principal must have expectations of success for all students, and a principal must take the lead in instructional planning." These suggestions, given to a principal with little formal training in reading, are not very helpful as he struggles to build a strong reading program.

Solution One

The role of the building principal in the reading program is complex and many faceted. However, it is necessary to understand some of the complexities if we are to help principals. Reading researchers need to join with other educational professionals to study specific ways in which the behavior of the principal affects the reading program. Carefully planned, long term ethnographic studies of the role of the building principal in developing and maintaining an exemplary reading program are needed. Factors which both encourage and inhibit principal-staff interaction in developing effective reading curriculums need to be discovered.

Problem Two

Reading professionals need to take a more active role in the formal preparation of elementary principals. In many instances the reading professionals have considered the training of administrators to be "outside of their province" while they concentrated on the teaching/learning environment within the classroom and the training of classroom teachers. Unfortunately, preparation programs for most administrators emphasize school law, school plants and facilities, school finances, and management techniques, and very little emphasis is placed upon instructional methodology and curriculum development. Standards for certifying administrators in most states do not stress reading.

The certification standards for principals were obtained from 47 states and carefully analyzed for this paper. There was no state that specifically listed courses in reading as a requisite part of the certification standards for principals. Many of the state certification officers stated, however, that a pre-requisite to administrative certification was either elementary or secondary education certification and that at least one course in reading was required for this certification. Therefore, they assumed that a reading course would have been part of an administrator's training.

All states, however, required proof of specific competencies in school finance, school law, and administrative theory before administrative certification would be given. Thirty-three states required a course in special education. Courses in specific curriculum subjects were usually grouped, and courses or competencies were required in certain areas. It is possible to fulfill the competency requirement in curriculum areas with a combination of courses in a variety of subjects or two or three courses in one subject. Reading courses could be taken as part of the curriculum area, but they would not have to be. No state specifically required a course in reading methods for certification as a principal. Therefore, if a candidate applying for administrative certification had received elementary or secondary certification before reading courses were mandatory, it is very possible that a candidate could receive administrative certification without taking any courses in reading. A strong special education lobbying effort made courses in special education mandatory for administrative certification. Perhaps a lobbying effort by reading professionals is needed to make courses in reading mandatory for elementary principals.

After the analysis of certification standards was complete, 149 elementary building principals in three states were interviewed by telephone. Each principal was asked the number of reading courses which he/she had taken. The sample contained

118 men and 31 women. Of the 149 subjects, 32 principals reported having had no course in reading at either the undergraduate or graduate level. Of the 149 subjects, 74 principals reported having had one undergraduate or graduate reading course, and 31 had two or three formal courses in reading. Twelve principals reported that they had taken four or more reading courses. Of these twelve principals, nine were female. The majority of the 32 elementary principals who reported taking no reading courses were secondary school teachers who were serving as elementary school principals.

It is not surprising that building principals, two thirds of whom had none or only one course in reading, did not feel qualified to show strong leadership in the area of reading.

Solution Two

Reading professionals need to become actively involved in setting up certification standards which would make reading mandatory for all elementary principals. Principals will choose the tasks which they feel most comfortable doing when they define their roles; principals with training in reading and language arts programs will be more apt to show leadership in the reading program.

Problem Three

Building principals are generally reluctant to become highly involved with reading program leadership because they feel inadequately trained to do so. In the few instances when they do become involved, it is generally to implement one or two strategies which they feel work well. However, in very few instances does the elementary school principal assume a highly visible and responsible role in the development of a balanced reading program without special encouragement. It is the duty of the reading professionals to help principals develop these sets of leadership skills in reading.

At this point, however, reading educators are not reaching the building principals through their courses. For this study the enrollment in the reading courses of three universities was surveyed to find out whether these reading departments were serving administrators or administration trainees. In the course which dealt with the administration and supervision of reading programs during one summer school semester, 56 students were enrolled in the courses. Of these 56 students, 5 students were principals and 2 students were working on administration certification. In the graduate developmental reading course 105 students were enrolled. Of these 105 students, two were principals and three students were working on administration certification. In the remedial reading courses, 62 students were enrolled. Of these students, none were principals and two were working on administrative certification. Thus existing reading programs are generally not reaching many elementary principals.

Solution Three

Principals need to be encouraged to take introductory reading courses to learn the basic fundamentals of reading. College and universities need to do active recruiting among principals and encourage them to attend reading classes.

However, colleges also need to help principals develop and maintain an excellent school wide reading program. A careful perusal of many course syllabi in reading departments shows that little attention is given to this aspect of developmental reading. A course designed specifically for building principals could have a two-pronged focus; the first part could stress fundamental reading knowledge, the second part could stress the role of the principal in building a strong reading program. This special course for principals also needs to be marketed carefully so that building principals realize that these courses could help them. And finally, reading professionals should encourage educators with strong curricular knowledge to enter administration.

Summary

This problems court addressed the major issue of the lack of participation of many reading educators in the development of effective elementary school principals. The major problems discussed concerned the lack of research in the role of the principal in the reading program, the lack of reading professionals' participation in the development of licensing and certification standards, and, the lack of courses taught specifically for administrators of reading programs. The solutions suggested include the need for careful research into the role of the principal in the reading program, the active lobbying for higher certification standards and required reading courses for elementary principals, and finally, the development of courses which will show principals how to develop and evaluate strong reading programs.

REFERENCES

Aspy, D. N., & Roebuck, F. N. (1977). *Kids don't learn from people they don't like*. The report of the National Consortium on Humanizing Education. Amherst, MA: Human Resource Development Press.

Georgiades, W. (1985). The Principal as an Instructional Leader: Myth or reality? *Educational Considerations*, *12*, 12-16.

Georgiades, W. (1978). *How Good is Your School*? Report of the Model Schools Project. New York: Danforth Foundation.

Spencer, H. W. (1980). *Of Projects and Principals*. Reston, VA: Association of Teacher Educators.

Weber, G. (1971). Inner city children can be taught to read: Four successful schools. Washington, D.C.: Council for Basic Education.

STUDENT INTERVIEWS AS A LEGITIMATE EVALUATION TOOL FOR ASSESSING SCHOOL READING PROGRAMS

JUDITH K. CASSADY
LAWRENCE G. ERICKSON
KATHLEEN GAARE-WEISE
DEBORAH J. ROGERS

Southern Illinois University

Currently, much is being written about how to evaluate the effectiveness of instructional programs and materials. Student interviews can be used as legitimate and valuable tools in assessing school reading programs. In fact, student interviews are a unique way of evaluating the effectiveness of instruction from the point of view of students.

Many types of evaluation tools have been suggested for reading programs and instructional practices ranging from self-appraisal to reassessing one's thinking about the teaching act. These methods suggest a kind of objective introspection in self-evaluation, which most of us know is a "tall order." Not unlike the Stanislovski method of acting, a teacher must be able to "step outside" himself or herself and, in a sense, be able to witness the total effect of a chosen plan for delivery of instruction. Even then it is difficult to know how one is being perceived by "the audience" since we seldom hear applause nor tolerate the throwing of ripe fruit in the classroom. In fact, learners are trained early not to openly express their feelings (nor to have any negative ones!) about the kind of instruction they receive. Furthermore, children are often unable to distinguish between "good" and "poor" textbooks or classroom practices even when directly asked.

The idea of checking student perceptions of texts and what and how they are being taught, caused us to approach student interviews as an action research project. We viewed this as formative evaluation -- a means of assessing on-going instructional practices -- which might help teachers become aware of the impact of their instruction. This procedure might be viewed as a metacognitive/metapedogogical process.

The concept of interviewing as a measurement tool in evaluating school programs is not new (Cannell & Kahn, 1968; Farley, 1981) and is actually quite simple and direct. But the construction and effective utilization of results from a good questionnaire are often an eye-opening experience if cognitive and affective dimensions are considered.

Developing Interview Questions to Generate Thought and Action

To begin this process, a brief but thorough set of questions which "gets at the heart" of a program must be developed. Based on our experience, it was obvious that questions to measure student perceptions can lead teachers to "step beyond" themselves in order to understand what they are doing as they are doing it. Interview questions based on "What do I want students to think, feel, remember or be able to do?" "Should I do this?" "How do I do it?" give teachers valuable feedback that can encourage improved teaching.

In addition, it is equally important that teachers themselves give anonymous responses to a questionnaire parallel to the one which the students use in order to provide a means of comparison for analyzing results later on.

Reliability and Validity of Student Interviews

Student interviews are plagued with some problems of reliability and validity. These problems include: (1) student understanding (or misunderstanding) of the purpose or role of the interview and interviewer; (2) student motivation to under or over report their perceptions; (3) student forgetting, repressing, or altering responses due to ego enhancing or threatening factors; and (4) biases due to interviewer and/or respondent roles and backgrounds. These factors are especially crucial if the interviewers are adults (teachers) and the respondents are children (students).

The reliability and validity of student interviews can be enhanced by paying attention to sampling, question construction, interview techniques, and reporting activities. For example, student interviews might be used as one source of data included with other data from questionnaires, observations, and tests. Items included in the student interview might tap the same instructional aspect across different data sources so that teacher perceptions can be compared to student, parent, and administrator views. A mixture of open and closed items is important, and asking students to report what they did within the last day or two in school helps to improve their accuracy. Above all, anonymity and the assurance of the confidentiality of responses is essential to the validity of the process. Responses should be reported only in summative terms, means or percentages to avoid responses that could be linked to individuals.

Use of the Results

The results of student interviews and teacher questionnaires should be tallied, averaged and disbursed in some understandable way. The purpose of interviewing and having the teachers question themselves is to provide avenues of comparison and/or contrast between the opinions of the students and those of the teachers.

This type of thought and action generating survey is intended to be a help to the school, and the results belong to the teachers and students. The results can be used to stimulate discussion of what, if any, areas might be improved and how. But the decision to try to change any practices or materials rests with the participants. It must be stressed here that the results also can be used to show where strengths exist, because it is sometimes difficult to see what is working.

The following questionnaires are examples which were written and administered by the authors for examining the effects of reading instruction in the content areas in grades 4, 5 and 6 in one mid-western community. Six classroom teachers in the school responded to the teacher questionnaire prior to the student interviews so that they would be aware of the types of questions their students would be asked. They were then asked to send two average, two above average, and two below average students to the cafeteria to be interviewed by the authors. Students were told that the interviewers were from the university, that they were interested in the kinds of things most students do in their classrooms, and that they were most interested in the opinions students held regarding what they did in the classroom.

This example represents an artificial situation in that the teachers did not write the questions themselves, and the teachers didn't do their own interviews. However, it was hoped that such an example would give them a better understanding of the process which they may subsequently choose to replicate with their own adaptations.

On this questionnaire (See Figure 1), teachers indicated that they felt there was ample time to read textbooks in class (Item 3); that maps, graphs, and charts are explained (Item 8); that, while study guides are provided in science and social studies (Item 9), they are not provided in math; and that students are told of the outcome for lessons (Item 10). For math, Item 11 (materials for different abilities) was ranked high reflecting the ability grouping that is used in math. The items teachers rated lowest were: their use of audio visuals (Item 13); their communication with others on what they do to help poor readers (Item 15); the provision of remedial help for students having difficulty in content areas (Item 16); and their providing pre-reading activities before students read (Item 2). They also rated Items 1 and 9 lowest in math. This appears to reflect the teachers' feeling that, since students are ability grouped in math, there is less need to provide help for poor readers, and study guides are not needed to read math texts.

Students interviewed on this same occasion (See Figure 2)

FIGURE 1: Teacher Questionnaire

What grade level do you teach? What subject areas?

Explanation: In almost every classroom there are disabled readers who might benefit if they receive help in comprehending science, social studies and math textbooks. Please respond to the following questions as they relate to the students in your class(es). Rate each question in each content area using the following scale:

1 - Never **2 - Seldom** **3 - Often** **4 - Always**

Sample Item:

	Science	Soc. Studies	Math
Disabled readers use the same textbook as the average and better readers in my class in . . .	1 2 3 4	1 2 3 4	1 2 3 4

(MEAN responses reported in parentheses above selections)

	Science	**Soc. Studies**	**Math**
1. There is a need to provide help for poor readers in . . .	(3.17) 1 2 3 4	(3.33) 1 2 3 4	(2.33) 1 2 3 4
2. I provide pre-reading activities before students read their assigned textbooks in . . .	(2.5) 1 2 3 4	(2.67) 1 2 3 4	(2.67) 1 2 3 4
3. Students have ample time in class to read their textbooks in . . .	(3.83) 1 2 3 4	(3.83) 1 2 3 4	(3.67) 1 2 3 4
4. Students are taught how to determine the main idea of passages in their textbooks for . . .	(3.17) 1 2 3 4	(3.0) 1 2 3 4	(2.83) 1 2 3 4
5. Students are taught how to determine word meanings from the context of the passage in . . .	(3.0) 1 2 3 4	(3.17) 1 2 3 4	(2.5) 1 2 3 4
6. Students are taught how to skim passages in their textbooks to find the general idea during . . .	(3.33) 1 2 3 4	(3.33) 1 2 3 4	(2.67) 1 2 3 4
7. Students are taught to scan their reading assignments to locate specific facts in . . .	(3.33) 1 2 3 4	(3.33) 1 2 3 4	(3.0) 1 2 3 4

8. I demonstrate how to read maps, graphs, charts, and other illustrations before my students read their assignments in . . .	(3.67) 1 2 3 4	(3.67) 1 2 3 4	(3.83) 1 2 3 4
9. I provide study guides for my students in . . .	(3.67) 1 2 3 4	(3.67) 1 2 3 4	(2.33) 1 2 3 4
10. I tell my students what the expected outcomes and/or the objectives of each lesson or unit are in . . .	(3.67) 1 2 3 4	(3.67) 1 2 3 4	(3.83) 1 2 3 4
11. I provide a variety of materials for students of differing abilities in . . .	(3.0) 1 2 3 4	(3.5) 1 2 3 4	(3.33) 1 2 3 4
12. I point out new words and/or concepts before asking students to read their assignment in . . .	(3.5) 1 2 3 4	(3.5) 1 2 3 4	(3.33) 1 2 3 4
13. I use A.V. (films, records, TV) to enhance concepts presented in . . .	(2.67) 1 2 3 4	(2.67) 1 2 3 4	(2.33) 1 2 3 4
14. Students who achieve the best reading grades also achieve the best grades in . . .	(3.0) 1 2 3 4	(3.17) 1 2 3 4	(2.83) 1 2 3 4
15. I communicate with other faculty what I am doing to help poor readers understand the content of . . .	(2.17) 1 2 3 4	(2.33) 1 2 3 4	(2.5) 1 2 3 4
16. Additional and/or remedial help is provided to students who are having difficulty in . . .	(2.33) 1 2 3 4	(2.33) 1 2 3 4	(2.83) 1 2 3 4

In a sentence or two, tell why you like or dislike the textbook/materials you have for science.

—for social studies.

--for math.

FIGURE 2: Student Interview

Suggestion: Put the student at ease by asking his/her first name, what grade s/he is in, and whether s/he likes school this year. Explain that you want to ask some questions and that the teacher will be told only how most people answered.

(easy 78%) (hard 8%)
1. Is your science book easy or hard to read? (mixed 14%)

   How about your social studies book? (easy 61%) (hard 25%) (mixed 14%)

   Are the directions or the story problems in your math textbook easy or hard to read? (easy 47%) (hard 31%) (mixed 22%)

2. Does your teacher help you with these books by telling you about the new words, the important ideas, or by discussing what you will be reading about before you read? (yes 75%) (no 25%)

3. Do you have time in class to read all of your assignments in social studies?

Never	Seldom	Often	Always
(0)	(7%)	(36%)	(57%)

   How about science assignments?

Never	Seldom	Often	Always
(0)	(7%)	(36%)	(57%)

4. Does your teacher show you how to read the maps, graphs, and charts that are in your textbooks?

Never	Seldom	Often	Always
(6%)	(19%)	(33%)	(42%)

5. Does your teacher remind you to look carefully at all the pictures in your textbook before you read your assignment?

Never	Seldom	Often	Always
(22%)	(39%)	(25%)	(14%)

6. Show me how you skim this page to find the general idea? (Use a copy of a page from the student's social studies text.)

Can skim well	Has some idea	Has no idea how to skim
(22%)	(64%)	(19%)

7. Show me how you would figure out the meaning of this word if you had never seen it before.

Can use context clues	Has some idea	Has no idea
(50%)	(36%)	(14%)

8. Show me how you would determine the main idea of this paragraph in your social studies book?

	Can determine the main idea of a passage	
(56%)	(44%)	(0)
Has the general idea of how to		Has no idea how to

indicated that science books were easier than social studies, and the math text was the hardest. It is interesting to compare teacher and student perceptions on ease of texts. For example, teachers comments showed that they liked the math text with comments like "excellent," "good," and "fine." Teachers also liked the science text with comments such as "interesting," "difficult," "clear," and "words highlighted." However, they described the social studies textbook as "boring," "too general," "not organized," "not detailed," "lacking coverage." It may be worth discussing these questions: What does it mean when the book most liked by the teachers (math) is the one rated most difficult by the students? Or what does it mean when the book rated easiest by the students (science) is also liked by the teachers?

A majority of students (75%) said that teachers helped before assignments were read (Item 2). However, teachers rated their pre-reading teaching rather low (Item 2). Perhaps teachers are doing more than they realize to help students before they read. Or another explanation might be a lack of a clear definition of pre-reading activities. Adequacy of help was also not evaluated by students in their responses.

Both teachers and students agreed that time is provided to read assignments in class. Teachers ranked time to read highest (Item 3), and 93% of the students said they had time in class (often or always) to read assignments.

A majority of students (75%) said that teachers showed them how to read maps, graphs, and tables. However, 61% said teachers seldom or never reminded them to look carefully at all the pictures before reading. Perhaps students (or teacher) perceived "pictures" as different from maps, graphs, and tables.

A majority of students (86%) indicated some grasp of what it meant to skim for the main idea of an article or could skim to find the general idea. Likewise, 86% had some grasp of, or could use, context clues to figure out word meanings. And 100% either had a general idea of, or could determine, main ideas of paragraphs in social studies texts.

There was high agreement between teachers and students on three items regarding:

1. Time - both teachers and students said there was ample time to read assignments in class;
2. Maps, Graphs, New Words - both teachers and students said students were helped with these before they read;
3. Skimming, Main Idea - both teachers and students said this was being done successfully.

There was less agreement on two points:

1. Pictures - Teachers ranked illustrations second from the highest, while 61% of the students said teachers "seldom" or "never" tell them to look at all pictures prior to reading;
2. Pre-reading Activities - Teachers ranked pre-reading activities low, while 75% of the students said teachers gave them pre-reading help.

These results suggest that pre-reading activities could be discussed, clarified, and perhaps examined more closely. The reading research is quite clear on the value of pre-reading preparatory activities. That is, students learn better when given help and direction prior to reading the text. On one hand, teachers said they did not do pre-reading activities; on the other, students indicated teachers did some. Other questions to explore are: What are valid pre-reading activities? Are teachers actually doing them?

The results of this student interview also suggest that reading in math is different from reading science and social studies books. Perhaps one thing learned from this survey that difficulties may be encountered if math texts are included with science and social studies texts for evaluation.

Some Side-Effects

Questions to be asked of students should cause teachers (or those developing the questions, whoever they may be) to become more aware, more analytical, about their intentions and their practices. In addition, the interview questions tend to "let students in on" what they can expect from instruction.

Questions which elicit student perceptions of what's good for them and how they like what they are doing in class often accentuate student awareness of instructional practices, materials, and their own skills which they might not have previously considered. In essence, such an interview, especially when there is respect for the interviewer, provides a student with a type of experience in metacognition. Once asked, he may continue to ask himself, perhaps in a subconscious way, if what he is doing is "effective" for him -- if this is the way he learns.

In Summary

In order to avoid some of the pitfalls inherent in the student interview process, the following "helpful tips" are offered:

1. Use single-idea or one issue questions.
2. Role play with "If I were a teacher....what should I tell kids about..." or "...how would you explain..." to help focus responses.
3. Use neutral questions to avoid biasing responses.
4. Use introductory and transition phrases.
5. Give feedback after responses to confirm with student that you are not misinterpreting responses.
6. If possible, train student interviewers to avoid role bias effects.

The use of student interviews can be a valuable tool for assessing student perceptions of your program. And after all, no matter how well one follows recommended practices even of the experts, the perception of the student is "what counts in the final measure." Moreover, if students have not been brought into awareness of the quality of materials or instructional practices, they are probably also unaware of their own skills or how they acquire them (or why they don't).

Student interview strategies are not difficult to develop or administer as long as advanced thought and preparation is adequate. The results can be useful and rewarding. A few words of advice:

1. Examine some samples of student interviews and consider results;
2. Be aware of the merits and pitfalls of using student interviews for evaluation; and
3. Consider using student interviews as valid sources of data in program evaluation and improvement efforts.

REFERENCES

Cannell, C. F., & Kahn, R. L. (1968). Interviewing. In G. Lindzu & E. Aronson (Eds.), Research methods: The handbook of social psychology, (Vol. 2). Reading, MA: Addison-Wesley.

This is a seminal reference for interviewing. Seventy pages are devoted to a thorough discussion of the interview as a measurement tool.

Farley, J. M. (1981). Student interviews as an evaluation tool. Educational Leadership, 39, 185-186.

This short article explains how to interview students. It is worth noting that this entire December, 1981, issue of Educational Leadership is devoted to evaluating school programs.

EVALUATION OF ADULT ILLITERATES ORAL LANGUAGE ASSOCIATIONS WITH THE PARADIGMATIC SYNTAGMATIC INVENTORY

JAMES A. DINNAN
University of Georgia

L. VERN PULLING
Literacy Action, Atlanta, Georgia

The authors view language/thought as one entity. Language is the outward expression of a thought through the rudiments of speech. The inward expression of the perception of the world about a human is a thought. An organized development of processing input stimulus enables the individual to retrieve and associate past experiences with the present and deal with the future. When one becomes a member of a language community, certain common relationships are expected in the exchange. If these relationships are not established communication will not take place, or if only some of the relationships are held in common, a poor form of communication will exist. Humans can perform as robots or computers. Put the information in, push a button, it comes out untouched, without any human intellectual processing. In order for one to "know" something, one must use a contrast because it is at that point a relationship is established. Hot is only established through cold, or cool, or warm etc., as a viable piece of communication. Both parties of the language community agree on the basis of establishing the conceptual framework and expect it to hold true within its parameter.

Thus paradigmatic type language/thought associations have been the expected associations when a subject is asked to respond to a given stimulus. (hot - "cold", in - "out", over - "under", man - "woman", is - "was", long - "short", none - "all".) Data collected over a wide range of areas has indicated that subjects who perform well on the tasks produce paradigmatic type response. Those who generally fall below expectations use a syntagmatic type association, which although meaningful to the individual, falls outside of the parameters of the common relationships used by the language community. (hot - "to trot", in - "bed", over - "the rainbow".) Table 1 presents the Oral P/S Inventory, an instrument for assessing paradigmatic/syntagmatic association.

Research into oral language associations, as a reflection of thought patterns, has been conducted with school children as subjects on a broad scale, Bigaj, J. J. (1977), Entwisle, D. R. (1966), Dinnan, J. A. (1980), Cartelli, L. M. (1980). However, the research of adult performance has been very limited, Katchen, (1980), Dinnan, (1971), Crable, (1975).

TABLE 1

ORAL P/S INVENTORY

1. car ____________
2. is ____________
3. on ____________
4. none ____________
5. long ____________
6. father ____________
7. only ____________
8. have ____________
9. over ____________
10. little ____________
11. one ____________
12. hat ____________
13. some ____________
14. top ____________
15. now ____________
16. are ____________
17. good ____________
18. any ____________
19. shoe ____________
20. front ____________
21. this ____________
22. be ____________
23. in ____________
24. all ____________
25. hand ____________
26. old ____________
27. many ____________
28. to ____________
29. bread ____________
30. year ____________

T	S	A	Q	M
2 8 15	3 9 14	4 7 11	5 10 13	1 6 12
16 22 30	20 23 28	18 21 24	17 26 27	19 25 29

Name ________________________________ Date ______________

Section ___________________

Total Score Pre P/S ________________ Date ________________

Total Score Final P/S ______________ Date ________________

Instructions P/S Test is to be given orally, in complete isolation from all other students. No encouragement or examples to be given at any time. Directions Give me the first word you think of when I say these words. (Allow 3 seconds for answer - otherwise mark as a blank)

Literacy Action is a non-profit organization, located in Atlanta, Georgia, dedicated to training adults to read, write, and compute in a literate society. Upon entering the program each of the 155 persons who served as subjects for this study were given the Oral P/S Inventory and were asked to read a passage from the Gray Oral Reading Test (1967) paragraph #6. Placement into the program at Literacy Action is based upon the results of these tests plus the Adult Basic Learning Examination - ABLE (1967) and a writing sample.

The five programs at Literacy Action range from 24 weeks of instruction through 72 weeks. The classes meet for one and one half hours twice a week with an attendance record reported of over 80% for the entire program. A sample of the academic achievements required before entering the next, more advanced stage are shown for Levels 1, 2, and 6:

Level 1.

a. reorganize alphabet
b. recognize at least 50 group generated sight words (in contrast)
c. write personal sentences using sight words
d. write personal sentences with total comprehension
e. recognize the 110 Prime Frequency Words in contrast

Level 2.

a. all of the achievements listed for 1
b. organize Prime Frequency Words in catagories of time, space, amount, quality and matter
c. write contrasting sentences using Prime Frequency Words and sight words
d. recognize 200 new sight words in contrast
e. write sentences using all sight words
f. use dictionary
g. know and apply basic word recognition

Level 6.

a. All of the achievements listed for 1,2,3,4,5
b. read passages pertaining to different subject areas, classify data relative to human relations to supernatural, nature, society, self
c. locate information on a topic in a library
d. use locaters (table of contents, index) to find information in a book
e. know how to apply test taking skills
f. have independent study skills

The subjects ranged in age from 19 years to 60 years old. The results of the testing were split into two sections because the subjects in one section had not finished the first part of the program and were not yet retested with the Oral P/S Inventory.

Section A (N=88) had a mean Pretest score on the P/S Inventory of 20.454. This indicated that for subjects in Section A at least 1/3 of their responses show a syntagmatic type association was taking place with a core concept of basic communication, the range of paradigmatic associations was from zero to thirty. The average time of participation in the program was 2 months.

The performance of the subjects on passage #6 of the Gray Oral Reading Test was poor for a large majority. Only six subjects passed the test (they were placed in level six of the program). The other 82 subjects failed to meet the requirements designated for passing and were placed in level one of the program.

Participants in Section B, N=67, those with both pre/post Oral P/S Inventory, provided the following data which were subjected to an analysis.

TABLE 2

COMPARISON OF PRETEST AND POSTTEST SCORES
ON THE ORAL P/S INVENTORY FOR SUBJECTS
AFTER 10.6 MONTHS OF INSTRUCTION

$N = 67$

X = Pre test

$\Sigma X = 1460$

$\Sigma X^2 = 35066$

Y = Post test

$\Sigma Y = 1721$

$\Sigma Y^2 = 44467$

$\Sigma XY = 38189$

$\overline{X} = 21.471$

$O_x = 7.450$

$\overline{Y} = 25.309$

$O_y = 3.686$

(Pearson) $r_{xy} = + .6728$

These data indicate that training in oral language associations can be done with adults of any age. The higher Posttest scores were achieved after an average time spent in the program of 10.6 months.

Subjects in the B section had a mean error of 12.22 on the Gray Oral Reading Test, passage #6, with a range from 1 to 23 errors. Initial program placement ranged from level two through level five.

When an adult uses personal (Syntagmatic) associations to oral or written communication a major breakdown can and will destroy the processing of data. The Oral P/S Inventory seems to be a reliable tool for assessing such performance. The remediation for a problem can be based on responses to the instrument. Literacy Action instructors feel that plans which are direct and teachable can be designed using the core of the language, the 110 Prime Frequency Words that account for over 50% of language usage. These are used as a basis for the five training programs implemented by Literacy Action personnel. The classification of time, space, amount, quality, and matter provide a matrix for storing and retrieving data introduced as "contrast sets" to participants in these programs. The general expectations of communication require each member of the group to use a common association, not an exclusive personal one. These adults have provided evidence of a working intellect and are capable of processing data according to general forms. They seem to lack understanding and skill in the rules of organization rather than intellectual deficit. A program designed to enhance fundamental associations can and does provide such a framework and allows the individual to enter the reading and writing code. Further research needs to be planned and conducted to assess the gains made by clients who participate in this program to its terminal level. Comparisons could also be made with groups who employ other instruments for assessment and other strategies for remediation.

REFERENCES

Bigaj, J.J., Dinnan, J.A., and Crable, E. (1977). Word contrasts: Beginning readers' responses. Journal of Reading Improvement 14, 32-35.

Cartelli, L.M. (1980). Reading comprehension: A matter of referents. Academic Therapy 15, 421-430.

Crable, E. (1975). A comparison of thought processes as measured by paradigmatic associations, grade point average, graduate record examinations, and faculty ratings, of foreign and native-born graduate students at the University of Georgia. Unpublished Doctoral Dissertation, University of Georgia, Athens.

Dinnan, J., Bickley, R., and Williams, J. (1971). Syntagmatic oral responses compared with high and low verbal SAT scores of college freshmen. Twentieth Yearbook of National Reading Conference, 17-19.

Dinnan, J.A. (1980). "Comprehension of Prime Space-Time-Amount Relationships by Learning Disabled and Normal Students" (with K. Sullivan and E. Mary Ann Cartelli). Journal of Reading Improvement 17, 3.

Entwisle, D.R. (1966). Word associations of young children. Baltimore: John Hopkins University Press.

Katchen, L.C. (1980). A Study in the Effects of Paradigmatic Language Training and its Transfer to the Reading and Writing Performance of Adult Illiterates. Unpublished Doctoral Dissertation, University of Georgia, Athens.

NONLITERATE ADULTS' KNOWLEDGE OF AND ATTITUDES TOWARD MICROCOMPUTERS FOR READING INSTRUCTION

LOUISE D.N. WACHTER AND EUNICE N. ASKOV
The Pennsylvania State University

Use of computer technology for instruction is currently being explored and advocated as one possible solution to the problem of adult literacy (Fingeret, 1984; Nickse, 1983). For example, the Gannett Foundation sponsored an invitational conference during the Fall of 1985. The purpose of the conference, which was entitled "Adult Literacy and Computers," was to bring together computer users and developers as well as knowledgeable practitioners in order to brainstorm ways in which technology might be used to deliver instruction.

In spite of some encouraging results (Buckley & Johnston, 1983; Imel, 1983; Rachal, 1984; Wilson, 1984) from the use of microcomputers for instruction, little is known about how nonliterate adults feel about using a microcomputer in order to learn to read. The attitudes of participants in any activity are important to its success (Clement, 1981). It is also not known whether certain groups of nonliterate adults feel more positively inclined toward computer-assisted instruction (CAI). For instance, are men more positively inclined than women to using a microcomputer for instruction? Is there a differentiation for age groups? Is an individual's attitude toward reading affected? This study was designed in an attempt to answer these questions.

METHODOLOGY

Instrumentation

A twenty-seven item questionnaire was developed by the investigators to assess three aspects of microcomputers: (1) knowledge of microcomputers; (2) attitudes toward microcomputers in general; and (3) attitudes toward using a microcomputer for reading instruction. For example, the question, "Does a microcomputer need instructions in order to run?" was labeled as knowledge of microcomputers; the question, "Does the thought of using a microcomputer make you feel nervous?" was designated as attitudes toward microcomputers in general; whereas the question, "Would you read more often if you could use a microcomputer for reading?" was labeled as attitudes toward using a microcomputer for reading instruction. The response options were: Yes, No, and Unsure. The instrument has a theoretical range of scores from 27 to 81, and an actual range of scores from 11 to 30. Since positive responses were scored with a 3 (negative responses were scored with a 1), the higher the score, the more positive the attitudes. The scoring system was reversed on the negative items (#2, 5, 8, 10, 11, 12, 15, 17, 18, 21, 22, 23, 26). The questionnaire is shown in Table 2.

The questionnaire was developed through several iterations from an original pool of 100 investigator-generated items. The items were randomly arranged. Of the 27 items, 14 items were positively worded (#1, 3, 4, 6, 7, 9, 13, 14, 16, 19, 20, 24, 25, 27) and 13 items were negatively worded (#2, 5, 8, 10, 11, 12, 15, 17, 18, 21, 22, 23, 26). In order to establish inter-rater reliability, three reading specialists served as judges. They worked independently in scoring the questions. The confidence ratings were (1) a judgement of how sure or unsure each judge was that an item was positively or negatively stated (r =.93) and (2) a judgement as to whether an item belonged to the assigned category (r = .96).

Because the questionnaire had not been tried-out with nonliterate adults, a pilot study was conducted in an Adult Literacy Program in central Pennsylvania in order to test the procedures and to assess subtest and total test reliabilities. Although the number of participants was small (N = 7), the age range and reading levels were similar to those in the present study. Four males and three females participated in the pilot study. Since the subtest reliabilities (1) Subscale 1, KNOW (r = .63); (2) Subscale 2, GLOBAL (r = .77); and (3) Subscale 3, READ (r = .89) and total test reliability (r = .86) were acceptable, only minor modifications were made in the wording of several items. Pilot testing also confirmed the appropriateness of using a three-point revised Likert-type scale as well as having the interviewers read the scale to the subjects. The revised questionnaire was used in the present study. It was administered during the Spring of 1985.

Subjects

The sample for this study consisted of 42 nonliterate adults placed on a waiting list for services in Adult Literacy Action, the field laboratory for The Pennsylvania State University Institute for the Study of Adult Literacy, Beaver Campus. The field laboratory is located in an economically depressed region of western Pennsylvania. The region surrounding the Penn State-Beaver Campus is comprised of both urban and rural populations. The subjects were randomly selected from the waiting list. There were 21 males and 21 females. Ages ranged from 19 to 59 years; their mean age was 27 years. Predominately, reading levels varied from second grade through sixth grade; however, four subjects had a reading level higher than sixth grade.

Procedure

Nonliterate adults participating in this study were asked to respond orally to the twenty-seven item questionnaire. In addition, biographical information was gathered pertaining to respondent characteristics such as, gender, age, and reading level. Information in regard to using a typewriter, playing video games, and the number of hours per week spent viewing television programs, was also collected. Two staff assistants to the Adult Literacy Action project director were selected and trained to administer the survey. The two assistants had previously worked with the adult clients in testing them and placing them on a waiting list for literacy services. Interviews were conducted by telephone. The subjects were told that the purpose of the questionnaire was to find out how they felt about microcompters. They were informed that their identity would remain anonymous, and that participation in the study was voluntary. In addition, subjects were told that there were no right or wrong answers. The participants responded to the questionnaire. Subjects were told that as each sentence was read, they were to think about it, then select one of the three possible response options: Yes, No, or Unsure. The participants had to remember the stem as well as the three response options. Each subject was allowed as much time as he/she needed to respond to the questionnaire. Approximately 15-20 minutes was needed to complete each survey. All responders provided the answer requested.

Data Analysis

Forty-two sets of data were collected, then scored and analyzed by the investigators. Two different analyses were performed.

Analysis 1. An estimate of the general performance of the scale was secured through computation of the mean, standard deviation, and alpha coefficients using the LIKERT program (Table 1). Although the scale was administered as a single scale, for scoring purposes, the items were regrouped into three different subscales: (1) knowledge of microcomputers (KNOW): items #9, 10, 11, 12, 16, 18, 19, 20; (2) attitudes toward microcomputers in general (GLOBAL): items #1, 2, 3, 4, 5, 6, 7, 8, 15; and (3) attitudes toward using microcomputers for reading instruction (READ); items #13, 14, 17, 21, 22, 23, 24, 25, 26, 27.

Analysis 2. A one-factor analysis of variance (ANOVA) was performed to see if the scale discriminated between gender, age, and/or reading level (independent variables) on the dependent variable, attitudes toward learning to read using a microcomputer (Subscale 3, READ).

RESULTS AND DISCUSSION

The intent of the investigators of this study was to determine nonliterate adults' knowledge of microcomputers as well as their attitudes toward microcomputers as a learning tool, and more specifically their attitudes toward using a microcomputer for reading instruction. The investigators also sought information as to whether gender, age, and/or reading level are differentially related to the subjects' attitudes toward using microcomputers for reading instruction. The results are interesting and perhaps not necessarily those that might be expected. Several points warrant discussion.

Subtest reliabilities were: microcomputer knowledge (r = .29), attitudes toward microcomputers in general (r = .11); and attitudes towards learning to read using a microcomputer (r = .79). Total scale reliability was established at .71.

TABLE 1

MEANS, STANDARD DEVIATIONS, AND
CORRELATIONS OF SUBSCALES

Subscales	N	Mean	SD	Correlation
Sub 1 (GLOBAL)	9	22.02	2.48	.288
Sub 2 (MICRO)	8	19.16	2.05	.106
Sub 3 (READ)	10	22.14	4.05	.788
Total Scale (TOTAL)	27	63.33	6.46	.710

Analysis 1. The low subtest reliability for microcomputer knowledge (r = .29) is likely attributable to the subjects' lack of experience with microcomputers. Since knowledge is not necessarily related to the measurement of attitudes, it was decided to eliminate this subtest.

The reliability for Subtest 2 was low (r = .11). This could be attributable to homogeneity of the participants on the construct being measured, in this case, attitudes toward microcomputers in general. The number of response options per question may have been a factor. Fewer response options affect reliability by restricting the range of possible scores. However, it may be difficult for nonliterate adults to have to keep the stem in mind as well as to have to remember three response options.

Subtest 3, *attitudes toward using a microcomputer for reading instruction* was the one of greatest interest to the investigators. Since the reliability for this subtest was satisfactory (r = .79), the decision was made to retain this subtest as the total test. Another reason for this decision was that feedback from the test administrators indicated that the questionnaire in its present form is too long for nonliterate adults.

Several items from Subtest 2, *attitudes toward microcomputers in general* will be retained because of the interesting results obtained to the questions: "Would knowing how to use a microcomputer make other people think more highly of you?" and "Would your family be proud of you if you learned how to use a microcomputer?"

Analysis 2. Attitudes toward using microcomputers for reading instruction did not vary with the independent variables (gender, age, reading level) on the dependent variable (attitudes toward using a microcomputer). An analysis of the data using a one-factor analysis of variance to determine significant differences between means yielded nonsignificant results. Therefore, there were no significant differences in the way males and females view microcomputers for reading instruction, nor were there significant differences in regard to age or reading level.

CONCLUSIONS AND IMPLICATIONS

The results show that the nonliterate adult students in this study do see a need for and a desire to learn about microcomputers for instruction. When the raw scores for each item on the twenty-seven item scale were converted to percentages (Table 2), it was determined that most subjects (74%) felt that the use of a microcomputer for instruction would make other people think more highly of them, and furthermore, that their families would be proud of them if they learned to use a microcomputer (96%). It is also worth noting that 76% of the participants expressed that they would like to have a job working with microcomputers. Nickse (1984) points out that the use of microcomputers for literacy instruction can prepare students for a job working with microcomputers.

TABLE 2

SUMMARY OF RESPONSES BY TOTAL GROUP

	QUESTIONNAIRE ITEMS	YES Raw Score/%	NO Raw Score/%	UNSURE Raw Score/%
1.	Is it important to learn how to use a microcomputer?	30 (71)	5 (12)	7 (17)
2.	Does the thought of using a microcomputer make you feel nervous?	28 (67)	13 (31)	1 (2)
3.	Would knowing how to use a microcomputer make other people think more highly of you?	31 (74)	5 (12)	6 (14)
4.	Would your family be proud of you if you learned how to use a microcomputer?	40 (96)	1 (2)	1 (2)
5.	Would you mind if a microcomputer told you that your answers were incorrect?	2 (5)	36 (85)	4 (10)
6.	Would you prefer to work alone at a computer instead of working with another person?	15 (36)	16 (38)	11 (26)
7.	Would you like to have a job working with microcomputers?	33 (78)	5 (12)	4 (10)
8.	Are you opposed to teaching elementary school students how to use a microcomputer?	3 (7)	32 (76)	7 (17)
9.	Does a microcomputer need instructions in order to run?	31 (74)	3 (7)	8 (19)
10.	Can microcomputers think and reason the way we do?	6 (14)	27 (65)	9 (21)
11.	Do microcomputers get bored?	1 (2)	33 (79)	8 (19)
12.	Do you have to understand everything about a microcomputer in order to use one?	21 (50)	16 (38)	5 (12)
13.	Could a microcomputer help you to learn to read better than you read now?	22 (52)	2 (5)	18 (43)
14.	Would you read more often if you could use a microcomputer for reading?	19 (45)	4 (10)	19 (45)

QUESTIONNAIRE ITEMS	YES	NO	UNSU
15. Would you be unwilling to enroll in a reading program that used microcomputers for instruction?	15 (36)	20 (47)	7 (1
16. Could a microcomputer help you to become a faster reader?	18 (43)	2 (5)	22 (5
17. Would reading from a computer seem difficult if pictures were not included with the sentences and stories?	11 (26)	19 (45)	12 (2
18. Can microcomputers read aloud to people?	14 (33)	18 (43)	10 (2
19. Could a microcomputer be used as a reading textbook?	22 (53)	1 (2)	19 (4
20. Could a microcomputer help you to learn to read a word that you do not already know?	18 (43)	5 (12)	19 (4
21. Would using a microcomputer for reading make you feel less confident?	7 (16)	25 (60)	10 (2
22. Would you prefer to have a person rather than a microcomputer teach you how to read?	27 (64)	5 (12)	10 (2
23. Would using a microcomputer tend to make reading more difficult?	5 (12)	11 (26)	26 (6
24. Do you think you would like using a microcomputer to learn how to read?	20 (48)	8 (19)	14 (3
25. If you could use a microcomputer for reading, would you like reading better than you do now?	15 (36)	6 (14)	21 (5
26. Would reading from a computer be boring?	7 (17)	17 (40)	18 (4
27. Is it a good idea to use microcomputers for reading instruction?	21 (50)	2 (5)	19 (4

The implications for those who are designing and/or using CAI for nonliterate adults are noteworthy. Based upon the results of the one-way analysis of variance, it appears that CAI may be appropriate for both men and women who are at various age groups and are functioning at various reading levels. Therefore, it seems that CAI reading instruction may be considered appropriate for adult literacy groups.

However, because of the important pedagogical effects of learner attitudes and the emergence of CAI as an important instructional delivery system, further study is warranted.

REFERENCES

Buckley, E. & Johnson, P. (1983). Pilot project in computer assisted instruction for adult basic education students. Albany: New York Education Department, Bureau of Community and Continuing Education Program Services.

Clement, F. J. (1981). Affective considerations in computer-based education. Educational Technology, 28-32.

Fingeret, A. (1984). Adult literacy education: current and future directions. Columbus, Ohio: ERIC Clearinghouse on Adult, Career, and Vocational Education.

Imel, S. (1983). The microcomputer: a tool for adult education. Adult Literacy and Basic Education, 7 (2), 109-115.

Nickse, R. S. (1983). The case for use of microcomputers in adult basic education. Quincy: Massachusetts Department of Education, Bureau of Student, Community, and Adult Services.

Rachal, J. (1984). Using CAI with ABE and GED students: a review of past and present programs. Presentation at the National Adult Education Conference, Louisville, KY.

Wilson, J. (1984). The ABLE Project: a model solution to illiteracy. Presentation at the National Education Conference, Louisville, KY.

THE DEVELOPMENT OF AN INSTRUMENT TO MEASURE TEACHER ATTITUDES TOWARD THE COMPUTER

SIDNEY E. BENTON
JUDITH L. LONG
North Georgia College

BOB W. JERROLDS
University of Georgia

During the last decade the use of the microcomputer in the classroom has increased as the ramifications of the Computer Age have reached the educational setting. A recent survey conducted by Market Data Retrieval (cited in Miller, 1985) revealed that while only 18% of the public schools in the United States had one or more microcomputers in 1981, more than 85% of the schools reported having acquired them by 1985. While these figures indicate the widespread availability of computers, the key to their effective utilization, as with any instructional material, rests with the teachers in whose classrooms the machines are placed.

Miller (1985) suggests that the effectiveness of the computer cannot be evaluated until the factors determining its acceptance or rejection for classroom learning are examined. Among the primary factors which may influence the extent to which teachers accept or reject utilization of computers in their classrooms are: (1) knowledge of the computer and (2) attitudes toward the computer. In an examination of those variables, Long and Benton (1984) found that "computer literate subjects displayed significantly more positive attitudes than did their non-computer literate counterparts" (p. 10). This finding would seem to suggest that because computer literate teachers are likely to have more positive attitudes toward the computer then they are likely to utilize computers in their classrooms.

If teachers' attitudes are indeed crucial determinants of their decisions as to whether or not to make use of computers, then it is important to study and understand their beliefs and attitudes in order to foster an environment that would maximize the potential for effective use of computers in instruction. A limited number of instruments examining attitudes toward the computer are available. The Computer Appreciator-Critic Attitude Scale (CACAS), a 40-item Likert-type instrument developed by Mathews and Wolfe (1983) and validated by Furr and Davis (1984), is one such evaluative instrument; however, the CACAS examines general attitudes toward computers and computer use rather than attitudes toward the educational uses of the computer. Likewise while the Reece and Gable (1982) ten-item Likert-type attitude instrument, Attitudes Toward Computers, is more school specific in that it was designed to measure student attitudes, the instrument still does not include items which would accurately reflect specific teacher attitudes toward the uses of the computer. Mitchell (1984) has developed a reading attitude instrument that has a section on student attitudes toward the computer, but again teacher attitudes cannot be extrapolated from the instrument. Therefore, it appears that if teacher beliefs and attitudes are to be examined, a more specific instrument to address those attitudes and beliefs is warranted. The purpose of this study was to develop such an instrument in order to measure teacher beliefs about and attitudes toward the educational uses of the computer.

Procedures

For the development of the instrument the investigators planned an empirical approach with two major phases. In the first phase, the goal was to obtain a large pool of items written by teachers. The items were then to be reviewed by the investigators for the purpose of categorizing the items and deleting or improving items. The items were then to be revised in terms of analyses and comments from a panel of judges. In the second major phase a principal components analysis was to be used to categorize the items. This second phase has not been completed and is not discussed in this report.

In the first phase of the study, the present investigators determined five rational areas that they thought reflected teachers' beliefs about and attitudes toward educational uses of the computer. These areas were instruction, remediation, enrichment, management systems (including testing, record keeping, and grading), and general attitudes.

Students in two graduate educational research classes (N=36) and one graduate/undergraduate class of reading education in the middle and secondary schools (N=12) were asked to develop items that reflected these five rational areas. They were to make one-half of their items reflect positive positions and one-half of their items reflect negative beliefs or attitudes. All of the students in these classes were either pre-service or inservice teachers. In addition, they had all received instruction in the development and the scoring of attitude scales. A total of 201 items were thus generated.

Two of the present investigators then sorted the items into the five categories and eliminated items that were considered to be duplicates. Items that the researchers deemed to be unclear or poorly written were also eliminated. In addition, minor word changes were made in some of the other items--thus, the original 201 items were narrowed to 85 items.

All ten students enrolled in the Reading Education Doctoral Seminar at the University of Georgia were asked to act as a panel of judges to answer the following questions with regard to each of the 85 remaining items: 1. Does the item belong in the category to which it is assigned? 2. Is the item essentially a duplication of another item? 3. Would the item be clearly answerable in the format of a semantic differential? 4. Does the item represent an attitude, or does it represent a belief? (Definitions of these terms appear later). 5. If answered as "Wholly agree," would the answer represent a positive attitude or belief or a negative attitude or belief? This panel of judges was told about how the instrument was developed, up to the point of their input, and what the authors hoped to have as a final instrument. They were asked to respond to all of the above questions and to make any comment on the instrument that they thought would be helpful to its developers. The responses of this panel were utilized to establish the content validity of the instrument.

In response to the question of whether the items were in the right categories, the majority of the panel recommended that only two items be moved to another category. One was moved as suggested. The other was deleted because it was deemed to be essentially the duplicate of another item in the category to which it was recommended that the item be moved. All other items were left in their original categories since no fewer than eight of the ten judges held that they were appropriately classified.

In response to the question of duplication of items, the majority of the panel recommended that four items be deleted because they were duplicates of other items. All four were deleted. The panel was in total agreement that all other items were not duplicates.

To check the question as to whether the statements generated by the item originators might have represented beliefs rather than attitudes, the present authors presented the panel with definitions of the two concepts. Attitude was defined by Wolman in his Dictionary of Behavioral Science (1973) as "A learned predisposition to react consistently in a given manner (either positively or negatively) to certain persons, objects or concepts. Attitudes have cognitive, affective and behavioral components" (p. 34).

English and English, in their Comprehensive Dictionary of Psychological and Psychoanalytical Terms (1958), defined belief as "An emotional acceptance of a proposition or doctrine upon what one implicitly considers adequate grounds. The grounds for belief, however, are often not examined, nor does the believer imply that others need have the same grounds. Beliefs have varying degrees of subjective certitude" (p. 64).

In not one case out of the total of 85 items was there 100% agreement as to whether the items represented an attitude or a belief. There were two cases in which one panelist said an item represented an attitude while eight said it represented a belief. There were ten items in which the split was two to eight. In all other cases the split was more nearly equal.

Since Wolman (1973) did not define belief and English and English (1958) did not define attitude, it is possible that the definitions were inadequate. No definitions were found that adequately explicated differences in the two concepts. Since the present writers, in their own minds, could not clearly distinguish between attitudes and beliefs nor even determine if such a distinction was necessary, it was decided to eliminate no item based on judgment that it represented a belief rather than an attitude.

The instrument was originally envisioned as a Likert-type scale. The designers of the instrument felt that a five-point scale was appropriate given the rather divergent feelings teachers seem to have about computer use in the classroom. Since the assumption of the equality of the intervals on scales having more than two intervals could not be adequately addressed otherwise, the instrument developers wanted to meliorate a possible problem by using a semantic differential format as utilized by the instrument at the end of this article. The panel of doctoral students were instructed as to the nature of a semantic differential and asked to indicate whether each item on the original instrument could be responded to adequately in a semantic differential format. The semantic differential format was judged to be inappropriate for three items by two out of the ten judges. In all other cases the judges were nine to one or 100% in agreement that the format was appropriate. Because a clear majority of the judges indicated that the semantic differential was appropriate for all items, no items were eliminated because of response format.

A further validity check was run on each item by having the panel of judges indicate whether the items represented a negative or a positive attitude or belief. It was decided that any item on which more than two judges disagreed with the key would be revised or eliminated. Using that criterion, only one item was eliminated, and one was revised. In the present instrument there are two items on which there was 80% agreement among the judges with the key. All others have stronger agreement.

The panel of judges was asked to make any comments on the instrument they thought might be helpful to the instrument developers. On the basis of these comments minor wording changes were made in some items. As a result of the analyses of the items by the panel of judges, the 85 items were narrowed to 79 items reflecting the five rational categories. The categories and the numbers of items in each category were: instruction-20, remediation-14, enrichment-11, management systems-23, and general-11. The 79 items were then taken out of their categories and randomly assigned to a position within the instrument. See a copy of the instrument at the end of this article.

The investigators are currently involved in the second phase of the development of the instrument. During this second phase plans are being implemented to eliminate, modify, and categorize the items as a result of a principal components analyses. It is anticipated that a later report will provide additional information about the development of this instrument.

REFERENCES

English, H. B., & English, A. C. (1958). A comprehensive dictionary of psychological and psychoanalytical terms. New York: David McKay.

Furr, J. D., & Davis, T. M. (1984, April). The validation of an instrument to measure attitudes toward computers. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans.

Long, J. L., & Benton, S. E. (1984, December). Survey and Analyses of Attitudes toward Use of the Computer. Paper presented at the annual meeting of the American Reading Forum, Orlando.

Mathews, W. M., & Wolfe, A. W. (1983, April). Measuring attitude toward computers: The Computer Appreciator - Critic Attitude Scale. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal. (ERIC Document Reproduction Service No. ED 231 886)

Miller, M. A. (1985). Computers in the classroom: Student attitudes toward an emerging technology in education. Unpublished master's thesis, North Georgia College, Dahlonega, GA.

Mitchell, R. M. (1985). The development and validation of a reading attitude assessment instrument for junior high school students. Unpublished doctoral dissertation, University of Georgia, Athens, GA.

Reece, M. J., & Gable, R. K. (1982). The development and validation of a measure of general attitudes toward computers. Educational and Psychological Measurement, 42, 913-916.

Wolman, B. B. (1973). Dictionary of behavioral science. New York: Nostrand Reinhold.

Teachers' Attitudes and Beliefs About Computers (TABAC)

Sidney E. Benton Judith L. Long
North Georgia College

Bob W. Jerrolds
University of Georgia

This scale has been designed to find out how you feel about computers. Consider each of the items carefully, and then record your response. There is no "correct" or "incorrect" answer for each item. In order to answer these items, you must determine how much you agree or disagree with the statement as it applies to you. For each item mark:

A (1) if you strongly agree with the item

B (2) if you agree with the item

C (3) if you are undecided about the item

D (4) if you disagree with the item

E (5) if you strongly disagree with the item

1. Teaching students to use computers in not worth the time.
2. The computer helps save time in preparing tests.
3. Computers cannot teach slow learning students.
4. Using the computer for enrichment frees the teacher to attend to students' needs.
5. Computer assisted instruction is boring and repetitious.
6. Computers simplify my record keeping.
7. Teachers would make fewer errors if they used computers for averaging grades.
8. Computers do not enhance instructional time.
9. Computers provide more individualization for students.
10. Computers are a waste of money.
11. Instruction by computers is too impersonal.
12. Students in need of remediation will only be confused by the technical aspects of a computer.
13. Enrichment using computers is impractical since there usually are not enough computers to go around.
14. I like to keep my records on a computer.
15. Computers are helpful in maintaining objectivity in grading.
16. Computer training for teachers takes more time than it is worth in the instruction of students.
17. Using a computer makes test grading much faster.
18. I feel very comfortable working with computers.
19. Computers are useful as instructional aids.
20. Computers are excellent tools for students who need extra work on skills because they give immediate feedback.
21. The computer software available is excellent for use with advanced students for enrichment.
22. It is too time consuming to maintain class records on a computer.
23. Using computers for grading consumes too much time.

24. The use of a computer for instructional purposes in the classroom is a threat to the teacher.

25. I would rather grade my tests by a computer than by hand.

26. There is insufficient use for computers in education as to justify their cost.

27. Nothing that I teach is appropriate for computer use in the classroom.

28. Computers are too impersonal to use with children with special needs.

29. Using the computer for student enrichment would improve student attitudes toward school.

30. It takes too much time to input information into computers for maintaining records and grades.

31. Keeping class grades on a computer facilitates analysis.

32. Using computers to administer tests is beneficial to both the student and the instructor.

33. Computers will replace teachers in the instruction of our children.

34. Students are enthusiastic toward computer-oriented instruction.

35. Students requiring remediation do not possess enough self-discipline to use the computer unassisted.

36. There is no value in using computers for enrichment.

37. Computers organize and maintain data faster, which allows teachers to do other aspects of their job.

38. Grading is done more efficiently with the computer.

39. Individualized testing should be done by computer to reflect individualized instruction.

40. Computers should be very helpful tools for teachers in any subject area.

41. Students seem motivated to learn when using the computer.

42. With low achievers, computer activities can be very motivational.

43. A computer could offer more varied enrichment activities for students in the classroom then they usually have.

44. User error can result in the loss of a lot of data at one time.

45. A computer would simplify grading by determining averages quickly and correctly.

46. Testing by computer heightens test anxiety.

47. Carefully prepared computer programs are excellent in assuring active learning.

48. I will enjoy teaching less if more computers are used in education.

49. Students end up playing with the computer rather than using it as a learning tool.

50. Computers are excellent for providing drill work for remediation.

51. Computers in the classroom should be used only as enrichment tools.

52. The problems in learning how to use computers outweigh any aid in record keeping.

53. I do not like to record grades on the computer.

54. Computer-assisted instruction frees the teacher to help other students who are having difficulty.

55. Slow learners can use the computer to work out problems and check their answers until the concept is clear to them.

56. Students resent computer-administered tests.

57. All students should have "hands on" experience with computers.

58. Computers interfere with students' instructional focus.

59. Computers are not an effective means of reinforcing skills.

60. Computer enrichment programs should be only for those who complete their work.

61. The computer is a nuisance in record keeping because too many people need access to it at the same time.

62. The computer has no instructional advantages and should no be used in educating children.

63. Teachers should not utilize computers in the classroom until there is more research to substantiate their worth.

64. Computer instruction takes more preparation than regular classroom instruction.

65. There is ample computer software available for remedial instruction.

66. Abstract thinking can be improved by using computers.

67. The computer is too complicated to be used for school record keeping.

68. Computers create discipline problems when they malfunction.

69. If a child has problems with learning, trying to use a computer is just one more thing to cause frustration in the child.

70. Students learn more when instruction is supplemented by a computer.

71. Computers serve as a good source of remediation activities.

72. Computers help the advanced students by challenging them with new ideas.

73. Valuable instruction time is lost in teaching student how to use computers.

74. All teachers should make use of computers in their classroom.

75. When used as an instructional tool in the classroom, computers enhance learning and improve teacher effectiveness.

76. Computer programs are just workbook pages on a machine.

77. There are many effective ways a computer can be used with remedial students.

78. Computers are only for the gifted and talented.

79. Computers are effective in communicating important concepts to students.

APPLICATION OF TEACHER EFFECTIVENESS RESEARCH FINDINGS

TIMOTHY R. BLAIR
Texas A&M University

JUNE THOMAS YOUNG
Alabama A&M University

The translation of research results into classroom reading programs has been a pervasive dilemma for teacher educators. The area of affecting teaching behavior is being given more importance in light of recent research on teacher effectiveness. This research has produced significant findings in the last decade (Heilman, Blair, and Rupley, 1986). Unfortunately, our knowledge regarding the fusion of these findings into practice has not been as significant. Good and Brophy (1977) conceptualized teaching to involve both a knowledge and performance dimension. Activities reported in this paper were an attempt to better understand the linkage between the two dimensions, i.e., ability to apply new knowledge in teaching situations.

Reasons for the slow incorporation of research findings into practice are many. An enormous quantity of literature documents the complexity of the change process. As related to changes in teaching practice, two major types of change, structural and attitudinal, must be acknowledged. For example, the changes many people recognize are the addition of materials and equipment to the classroom or school. This is only structural change--real change may be expected only when teacher attitudes are considered.

The most popular approach to effecting change in inservice programs is predicated on an organizational or bureaucratic model (Brubaker and Nelson, 1975). The inservice delivery system usually proceeds through the following steps: problem identification, consideration of possible solutions, decision on strategy or methods, implementation of decision, and the evaluation component. The missing factor that is not given primary consideration in this problem-solving model is teacher attitudes. Brubaker and Nelson direct attention to the importance of attitudes and the complexity of the change process:

> Some people make the mistake of viewing the educational change process as primarily a set of skills. Once these skills are properly

> exercised, the desired change or series of desired changes are set into motion. This position is based on a relatively simple and sequential idea of behavioral change: once the goals are set, the proper stimulus must be invoked on the selected target in order to achieve the desired response. This skill development position fails to take into account the fact that people are more than they appear to be. That is, others' perceptions of a person's behavior fail to adequately or validly take into account some of the following factors: attitudes, intentions, motives, and basic assumptions about reality (p. 63).

An accumulating body of literature suggests that effecting and sustaining change through the traditional bureaucratic model is extremely difficult and more attention should be paid to teacher attitudes and local characteristics (Berman and McLaughlin, 1978; Good, 1981, Mertens and Yarger, 1981; Samuels, 1981). Good noted that teachers expressing disagreement with new techniques tend not to introduce the suggested practices. In the Rand study by Berman and McLaughlin, 293 federal projects were evaluated regarding the establishment and maintenance of educational change. The authors found that successful projects were fragile and had difficulty in maintaining their success over a number of years. Also, successful projects achieved less than in the original sites. The Rand study highlighted the implementation stage as being the crucial stage in whether or not a proposed change would be successful or not.

In discussing the problem of implementing research results on effective teaching in the schools, McDonald (1980) stated: "I suggest to you that solving this problem requires translating results into practical activities. The next step is to do the developmental research to evaluate the effectiveness of these practical activities." McDonald goes on to say inservice developers should "study the teacher's perceptions, beliefs, and expectations" and design inservice programs that modify teaching styles rather than isolated teaching behaviors."

The inservice activities described here were an attempt to better understand the staff development process with respect to translating research findings related to the use of the direct instruction format into classroom practice. With student mastery of basic skills in reading the ultimate goal, there emerges overwhelming evidence from the research literature that classrooms should employ a direct instruction format. Summarizing the results from the Beginning Teacher Evaluation Study (BTES) on effective teaching practices, the Educational Testing Service (ETS) (1978) reported that "The critical elements in these effectiveness practices were direct instruction by the teacher and direct interaction between the teacher and the pupil" (p. 9). Similarly, the elements of direct instruction help to increase both student variables of opportunity to learn and the amount of academic learning time. The following description of direct instruction is provided by Rosenshine (1979), one of the original developers of the term:

> ...direct instruction refers to teaching focused on academic matters where goals are clear to students; time allocated for instruction is sufficient and continuous; content coverage is extensive; student performance is monitored; questions are at a low cognitive level and produce many correct responses; and feedback to students is immediate and academically oriented. In direct instruction, the teacher controls instructional goals, chooses material appropriate for the student's ability level, and paces the instructional episode. Interaction is characterized as structured, but not authoritarian; rather, learning takes place in a convivial academic atmosphere (p. 42).

Also, synthesizing results from recent experimental studies, Rosenshine (1983) identified the following six instructional "functions" as common components in describing effective instruction:

> Checking previous day's work, Presenting new content/skills, Initial student practice (and checking for understanding), Feedback and corrective (and reteaching, if necessary) Student independent practice, Weekly and monthly reviews

A change strategy focusing on providing options to teachers and a form of teacher self-reporting was employed in this inservice activity. Two groups of 20 elementary teachers (including one principal) each received two all day inservice sessions on teacher effectiveness findings with particular emphasis on the direct instruction approach. These before school sessions included both the traditional knowledge dissemination through lecture/discussion and demonstration of direct instruction lessons. Efforts were made to keep sessions informal in the interest of effecting a more personalized relationship between the teacher educators and the teachers. A follow-up activity suggested to participants was a teaching diary detailing direct instruction lessons completed in their classrooms. Support for using personal documents or journal writing as a legitimate method to study teacher thinking has been reviewed and discussed by Yinger and Clark (1985). Speaking to the efficacy of journal writing as a research tool, the authors concluded: "The danger of serious error resulting from use of journal entries as data seems small, while the promise of learning more about the psychology of teaching from the teacher's point of view looms large" (p. 28).

Specifically, two areas were to be highlighted in the diary: first, how teachers implemented the presentation or step-by-step explanation stage and second, how teachers implemented the guided or initial student practice.

Two group meetings were planned to discuss implementation of the direct instruction approach. The group sessions were designed to be nonevaluative and to have teachers become more conscious of what they were doing or not doing. It also provided a forum for sharing classroom practices documented in their diaries. Of the forty participants attending the initial two-day session, thirty-eight of them attended at least one of these voluntary follow up sessions, the largest attendance at a joint meeting in conjunction with the state reading conference. The majority of the participants also submitted a written copy of direct instruction lessons and/or details of changes implemented in the classroom as a response to the inservice activities.

Oral responses and written entries from teachers' journals revealed great differences in implementation. Some teachers detailed an extensive revision of their reading programs and lessons to more closely parallel the direct instruction format; others demonstrated or described lessons using games and activities of apparent many years usage.

While these activities are exploratory in nature, initial results are promising in understanding the complexity and individuality of change. The next logical step would be to systematically test the effectiveness of keeping a teaching diary in relation to actual classroom behavior and student achievement scores. With recent research affirming the importance of the teacher on students' reading achievement, it follows that strategies for inservice programs should focus on the unique styles and personalities of teachers. If reading programs are to change in our schools, so too teachers must change. Knowledge and transmission of characteristics of effective reading programs will not reach their fruition until innovative, teacher-centered inservice delivery systems are practiced by educators.

REFERENCES

Berman, P. and McLaughlin, M. (1978). Federal Programs Supporting Educational Change, Vol. VIII: Implementing and Sustaining Innovations (USOE). Santa Monica, Calif.: Rand Corp.

Brubaker, D.L. and Nelson, H., Jr. "Pitfalls in the Educational Change Process". Journal of Teacher Education, 26, 63-66.

Educational Testing Service, (1978). Focus: Learning To Read. Princeton, N.J.

Good, T.L. and Brophy, J.E. (1977). Educational Psychology A Realistic Approach. New York: Holt, Rinehart and Winston.

Good, T. "Teacher Expectations and Student Perceptions: A Decade of Research". Educational Readership 38, 415-422.

Heilman, A.W., Blair, T.R., and Rupley, W.H. (1986). Principles and Practices of Teaching Reading. Sixth Edition. Columbus: Charles E. Merrill Publishing Company.

McDonald, F.J. (1980). "Using Feedback to Change Teacher Behavior" In R. Weisman and P. Casini (Eds.), Three Views on Improving Basic Skills Instruction. Philadelphia: Research for Better Schools, Inc.

Mertens, S. and Yarger, S. (1981). Teacher Centers in Action: A Comprehensive Study of the Federally Funded Teacher Centers. Syracuse, N.Y.: Syracuse University and Syracuse Area Teacher Center.

Rosenshine, B. (1979). "Content, Time and Direct Instruction." In H. Wlaberg and P. Peterson (Eds.), Research on Teaching: Concepts, Findings and Implications. Berkeley: McCutchan Publishing Co.

Rosenshine, (1983). "Teaching Functions in Instructional Programs". The Elementary School Journal 83, 335-52.

Samuels, S.J. (1981). "Characteristics of Exemplary Reading Programs" In J.T. Guthrie (Ed.) Comprehension and Teaching: Research Reviews. Newark, Del.: International Reading Association.

Yinger, R.J. and Clark, C.M. (1985). Using Personal Documents to Study Teacher Thinking (Occasional Paper No. 84) East Lansing, Michigan, Institute for Research on Teaching, Michigan State University.

REACTION PAPER

Commentary on APPLICATIONS OF TEACHER EFFECTIVENESS: RESEARCH FINDINGS By Timothy R. Blair and June Thomas Young

Reactor: CAROL F. STICE
Tennessee State University

Blair and Young have addressed an issue in which all teacher educators are keenly interested, i.e. teacher effectiveness training. In short, Blair and Young have devised an in-service program for teachers that combines two major elements: (a) effective teaching plan consisting of a six step format for direct instruction adapted from Rosenshine (1983) and (b) a metapedogogical awareness component on the part of participants by having them keep a "teaching diary" over a two month period following the two day in-service. Blair and Young found that the teachers who participated in keeping the teaching diary and in follow-up discussions of their perceptions were more "susceptible to change" than those teachers who went through the in-service without the follow-up benefit of their own meta-cognitive/affective feedback.

The authors suggest that these findings occur because such meta-pedogogical feedback helps teachers become aware of their choices, allows them to feel in control of their teaching, and assumes that teachers are competent. Blair and Young believe learning and change occur more readily under such positive, supportive conditions.

The only difficulty I find with this approach is that it was developed to teach teachers how to teach for mastery of sub-skills - an approach which some reading educators believe establishes opposite conditions for children. That is, skills mastery teaching tends not to allow children many real choices, it does not encourage children to feel in control of their learning and it assumes learners know nothing or very little.

It seems to me that effective learning models are effective learning models whether they are applied to teacher learning or student learning. If such a meta-pedogogical approach to teacher in-service creates conditions for lasting change, and I have every reason to believe it does, I would prefer to see such an approach applied to content that promotes the same positive, supportive curricula for children; that is, teaching in such a way that children become aware of their choices, feel in control of their learning and are assumed to be competent.

The question becomes: If we have effective technique for creating change how then do we create change for the better? Perhaps an even more crucial issue is how to finally arrive at some real concensus among reading educators in light of what is now known about language learning and literacy.

READING EVALUATION IN THE CLASSROOM BROADENING ASSESSMENT AND INTERVENTION

LANCE M. GENTILE
University of North Carolina/Asheville

To date, much of what has been done to measure and remediate reading difficulties is based on standardized or informal skills tests designed according to the "scientific paradigm." Underlying this construct is the notion that if teachers identify one-to-one correspondents, i.e., skills deficits, an instructional "cure" can be found to relieve symptoms of reading difficulty. Silverman, Fite, and Mosher (1959), summarized this dilemma: "Considering reading difficulties as a symptom, if all efforts then are aimed at curing the specific symptom, we might expect these efforts to be of no more avail than painting away the spots to cure the measles" (p. 299). Perhaps, more than anything else, the reductionist view of reading may have contributed heavily to a dearth of research evidence demonstrating long-term gains and lasting effects of remedial reading instruction. Johnston (1985) noted:

> Past attempts at explanations of the differences between good and poor readers have tended to dwell on the minutiae of mental operations without considering either the psychological or social contexts in which they occur . . . Most current explanations of reading difficulties focus on the level of operations, devoid of context, goals, motives or history. While some work has focused on the context of reading failure rather than mental operations (McDermott, 1977; Mehan, 1979), there has been little effort to integrate these two dimensions. The consequent explanations of reading failure are sterile and have resulted in more or less terminal diagnoses of reading failure. Until we can integrate the depth of human feeling and thinking into our understanding of reading difficulties, we will have only a shadow of an explanation of the problem and ill-directed attempts at solutions. (pp. 174-175)

While some school administrators and state legislators continue to emphasize the use of "scientifically" developed reading skill tests for purposes of comparing groups of students and students within or among groups, this focus creates confusion and frustration among many teachers concerned with working to help students overcome their reading difficulties because:

1. It engenders the common misperception that because it is relatively easy to test and evaluate students' basic

reading skills deficits that this is more important than trying to understand, measure or infer complex social/ emotional variables that form the core of many students' struggle or failure in reading. Thus, for those not directly responsible for helping students surmount their reading difficulties, it appears that quantifiable information relative to grade level achievement or skills deficits is most significant. However, for classroom teachers this information provides very little useful data for purposes of helping students make a better adjustment to reading. For some, when it comes to measuring students' reading, it would appear "easily understood, workable falsehoods are more useful than complex, incomprehensible truths."

2. It creates the impression that the "whole" of reading is equal to the sum of the parts. Undeniably, skills are important in reading as is time-on-task skills practice but social and emotional adjustments determine how little or how much a student brings to and derives from reading (Dreikurs, 1954; Gentile & McMillan, 1981).

3. It causes them to delimit or ignore social and emotional issues when seeking explanations or solutions and designing instructional approaches and curricula to help remediate students' reading difficulties. And, according to Moffett (1985), "Then as now, if it's not tested it's not taught" (p. 52).

4. It is misleading because many students with reading difficulties can demonstrate a certain level of skills mastery (this may be slightly or greatly depressed compared to current age or grade level placement) (Lamb, 1985). In these cases, skills tests indicate the approximate level at which a student stopped or quit learning. But, significantly missing from these test results are the two most important pieces of information relevant to helping students learn to cope effectively with their reading difficulties, i.e., how hard did they try and what social/emotional (stress related) factors may have interfered with their performance on the test (Gentile, McMillan, & Swain, 1985).

Students' perceptions of their reading capabilities affect what they choose to do, how much effort they mobilize, how long they will persevere in the face of difficulties, their thought patterns, and the amount of stress they experience in taxing situations (Bandura, 1985). Successful attempts, therefore, to remediate students' reading difficulties must include not only skills development but also an emphasis on helping them acquire beliefs and convictions in their abilities to master these problems. There is a great difference between possessing reading skills and being able to apply or use them under challenging, difficult, or stressful circumstances. These situations require students with reading difficulties to develop self-regulatory skills in order to cope effectively. In the absence of these skills most resort to self-defeating stress reducing behaviors (Cannon, 1915, described these as "fight or flight" reactions) because they perceive reading as a threatening event. They envision a loss of self-esteem, failure, ridicule, incompetence, or rejection as a result of having to read. The following maladaptive stress reactions among these students to reading run a continuum from resistance to rage (fight) and avoidance to immobility or retreat (flight) (Gentile & McMillan, 1984; Swain, 1985). These apply to students who show no signs of neurological impairment or other physiological defects that would hinder normal development in reading.

Fight:

Resistance

1. Refusal to carry out reading assignments
   - Failure to complete reading tasks
   - Lack of perseverance; student refuses to read unless forced
   - Failure to see need to do anything students does not want to do; throws temper tantrums
   - Failure to make an honest effort to do well during reading even when closely supervised
   - Defies teacher's rules and authority
2. Refuses to follow directions that pertain to reading assignments
   - Does not attend to, remember, or follow directions concerning reading activities

Rage

3. Disruptive behavior during reading
   - Purposefully interferes with the teacher's instructional efforts
   - Talks without permission and interrupts other students or teacher frequently
   - Displays a marked hostility to reading, the reading teacher, or other students during reading or when asked to read
   - Frequent crying/whining or loud and raucous laughter during reading instruction
   - Other "acting out" behaviors during reading and reading instruction; extremely aggressive reactions.

Flight:

Avoidance

1. Avoidance of challenging reading tasks

   - Student shows a preference to be read to rather than make personal effort to read
   - Student continuously complains that reading is "hard" or "too hard"
   - Student avoids reading whenever possible, finds many reasons to become distracted.

Immobility

2. Student is reluctant to take risks in reading

   - Inordinate concern over being "right" and "reading correctly"
   - Hesitancy to read in front of others
   - Tendency to equate reading failure with rejection
   - Pattern of working slowly and procrastinating during reading
   - Indecisive, overly dependent behavior during reading or reading instruction

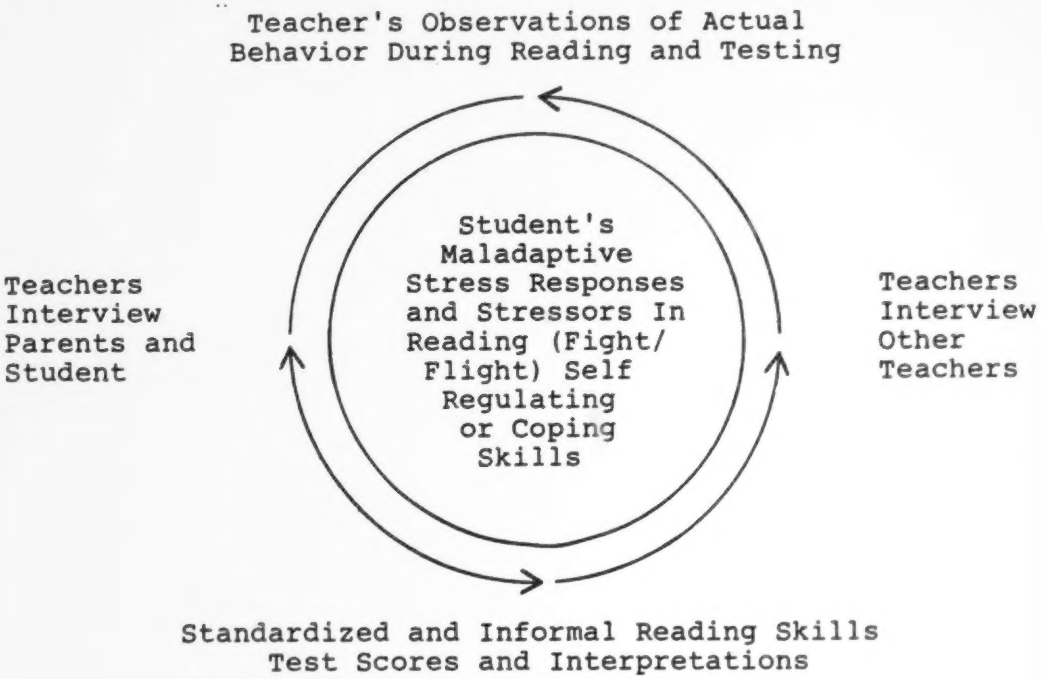
Retreat

3. Student withdraws from reading tasks that seem too challenging

   - Tendency to give up easily, become discouraged and quit. Does no independent reading.
   - Needs constant supervision and encouragement to complete reading tasks
   - Lacks confidence to complete reading assignments and reach goals
   - Appears apathetic, lethargic, and unwilling to try to overcome reading difficulties; is passive, unassertive, or immobilized during reading

If teachers' knowledge of students' reading difficulties and their ability to provide effective remedial instruction is to develop into understanding, it will not be the result of standardized or informal skills tests. In fact, focusing on the results of these tests, in terms of reading deficits, in order to develop remedial instruction, may create some deleterious side effects (Gentile, 1982). At best, skills deficits only partially describe students' reading difficulties. Of greater significance are the stress and stressors (Selye, 1976, spoke of situations, events, or people who produce stress reactions as stressors) students experience through early reading difficulty or failure These have the potential, in and of themselves, to disrupt all future experience with printed matter (Laurita, 1985).

The following figure presents visually the concepts discussed in this paper.


Teacher's Observations of Actual Behavior During Reading and Testing
Teachers Interview Parents and Student
Student's Maladaptive Stress Responses and Stressors In Reading (Fight/ Flight) Self Regulating or Coping Skills
Teachers Interview Other Teachers
Standardized and Informal Reading Skills Test Scores and Interpretations

References

Bandura, A. (1985). *Social learning theory*. Paper presented at the Carnegie Conference on Unhealthful Risk-Taking Behavior in Adolescence. San Francisco, CA.

Cannon, W. B. (1915). *Bodily changes in pain, hunger, fear and rage*. New York: Appleton-Century.

Dreikurs, R. (1954). Emotional predispositions to reading difficulties. *Archives of Pediatrics*, *71*, 340.

Gentile, L. M., & McMillan, M. M. (1981). Profiling problem readers: Diagnosis and prescription. *Academic Therapy*, *17*, 47-56.

Gentile, M. M., & McMillan, M. M. (1984). *Stress as a factor in reading difficulties*. Paper presented at Annual American Reading Forum Conference, Orlando, FL.

Gentile, L. M., McMillan, M. M., & Swain, C. (1985). Parents' identification of children's life crises: Stress as a factor in reading difficulties. In G. H. McNinch (Ed.), Reading research in 1984: Comprehension, computers, communication. Fifth Yearbook of the American Reading Forum (pp. 68-71). Carrollton, GA: Thomasson Printing and Office Equipment Company, Inc.

Johnson, P. H. (1985). Understanding reading disability. Harvard Educational Review, 55, 153-177.

Lamb, P. (1985). The symptoms of childhood depression as factors in children's reading difficulties. Doctoral dissertation, North Texas State University, Denton, TX.

Laurita, R. E. (1986). Child's reading ability linked to development of visual perception. Education Week, 5, 18.

McDermott, R. P. (1977). Social relations as contexts for learning. Harvard Educational Review, 47, 198-213.

Mehan, H. (1979). Learning lessons. Cambridge, MA: Harvard University Press.

Moffett, J. (1985). Hidden impediments to improving English teaching. Phi Delta Kappan, 67, 50-56.

Selye, H. (1976). The stress of life. New York: McGraw Hill.

Silverman, J. S., Fite, M. W., & Mosher, M. M. (1959) Clinical findings in reading disability children: Special cases of intellectual inhibition. American Journal of Ortho-psychiatry, 29, 298-314.

Swain, C. (1985). Stress as a factor in primary school-children's reading difficulties. Doctoral dissertation, North Texas State University, Denton, TX.

AN ANALYSIS OF LINGUISTIC AND COMPREHENSION STRATEGIES OF LOW ABILITY READERS

VICTORIA J. RISKO
Peabody College of Vanderbilt University

BACKGROUND

Mature oral reading behavior is often characterized as containing contextually constrained miscues with high use of self-corrections that are appropriate semantically and syntactically. The proficient reader is described as one who does not need to rely on all of the signals of written text to anticipate and confirm meaning (Goodman, 1965, 1967, 1969, 1973, 1985). Conversely, the poor reader has been characterized as a passive learner who is limited or inefficient in the use of reading strategies. When interpreting students' miscues during oral reading, Goodman (1973) indicated that while both good and poor readers use the same processes, the more proficient readers are more sensitive to semantic information and the less proficient readers use more grapho-phonic information and consequently lose more meaning. Comprehension is enhanced by increased attention to the semantic and syntactic information of text (Goodman, 1973; Goodman & Goodman, 1977; Smith, 1971).

Other reading researchers have indicated that proficient readers do not rely on contextual information more than graphic information, and that poor readers cannot be characterized as responding primarily to graphic cues (Allington & Strange, 1977; Biemiller, 1970). Weber (1970) proposed that children across all ability levels approach reading with similar linguistic abilities. Paris and Myers (1981) hypothesized that good readers may use the same strategies as poor readers but may use them more frequently and proficiently.

Previous research that has attempted to address the controversial issue of student use of strategies by evaluating and characterizing oral reading and comprehension behaviors of low ability readers has been limited in scope and application. First, as Leu (1982) indicated oral reading research has been flawed with several methodological problems (e.g., inadequate classification of miscues, lack of attention to the difficulty level of passages, lack of specificity regarding students' ability levels) that have limited the applicability of oral reading evaluation. Second, most studies have compared "good" and "poor" readers according to group trends rather than assess how students within ability levels may differ across assessment measures. Other studies that have assumed a single factor such as decoding or perception accounts for variance of performance between good and poor readers have obscured the heterogeneity of abilities within groups of readers (Singer, 1982).

The study presented here is part of a larger project that is being implemented to examine more precisely the oral reading behaviors of students classified as poor readers. The differential use of oral reading strategies by these students, using the factors of age and ability to describe subgroups, will be discussed.

METHOD

Subjects

Students (n=104) participating in this study were enrolled in grades 2 through 8 and were classified as having "reading problems" by their classroom teachers. They were attending an after school remedial reading program at the Family and Child Study Center of Peabody College of Vanderbilt University. About 90% of the students attended a public school and approximately 40% of the students were classified as economically disadvantaged based on parent salary and/or participation in a school lunch program. Almost all of the students received basal reading instruction that was supplemented by a basic skill, phonic-oriented program. There were seventy boys and thirty-four girls.

Assessment

This study provides a differential evaluation of readers' use of oral reading and comprehension strategies. In an attempt to accommodate methodological weaknesses of previous studies, consistent criteria were applied systematically to analyze all data. For this study, each student's oral reading comprehension was evaluated with a formal reading test, Gilmore Oral Reading Test (Gilmore & Gilmore, 1968), and a selected narrative passage from the Reading Miscue Inventory (Goodman & Burke, 1971). Error rates were established to be certain that each student read a passage appropriate for his/her difficulty level. Each student was evaluated individually. All oral reading miscues were scored systematically and in two phases to identify (1) use of graphic, phonic, and grammatical cues at the word level, and (2) use of language and meaning cues at the sentence and passage level. Retellings were analyzed holistically according to recall of text-explicit and text-implicit information. All data collected for the study were analyzed by two or more independent judges with reliability coefficients established at .90 or above.

Data analysis

For the data analysis, the students were divided into subgroups according to age and ability levels. The age categories and number of students within each were ages 7-8, twenty-two students, 9-11, sixty-nine students, and 12-14, thirteen students. The stanine scores of the oral reading accuracy subtest of the Gilmore Oral Reading Test were used to assign students to ability groups. The stanine clusters and number of students within each were stanines 1-3, forty-six students, stanines 4-6, fifty students, and stanines 7-9, eight students. Even though these students were reported to have reading problems in school, their performance on the Gilmore test, used here as one criterion of oral reading ability, indicated a wide range of abilities.

DATA SOURCE

The data were the scores obtained on the two measures of oral reading and reading comprehension. When percentages were used, tests for skewness and kurtosis revealed that all data points were well within acceptable limits of a normal distribution. A careful examination of relationships between and among miscues and comprehension performance scores was conducted to examine patterns and to make inferences about the readers' use of strategies. Individual profiles were compiled also.

RESULTS

The data were analyzed statistically using analysis of variance. Results indicated that low ability readers across age levels differed in their use of linguistic information sources. Older students (i.e., ages 9-11 and/or 12-14) were significantly different from the younger students (ages 7-8) in their higher use of grammatical, syntactic, and semantic information and self-corrections. These two groups of older students were also significantly different from the youngest group as evidenced by significantly fewer occurrences of meaning changes to the story. There was little difference across age levels on use of graphic and phonic information. (See Table 1.)

Across ability levels, students classified as low and average readers according to Gilmore's oral accuracy criterion, had a similar pattern in their use and frequency of linguistic strategies with the use of no one strategy differentiating group performance. Reliance on several linguistic strategies indicated a varied use of strategies for both groups. Students who scored higher on the Gilmore oral accuracy subtest tended to use the same stragegies as the other two groups, but they had higher scores on all areas assessed. On the use of grammatical and semantic strategies and self-corrects the students of higher ability were significantly different from the other groups. No single factor was identified to differentiate group performance. (See Table 2.)

Table 1. Means and standard deviations of miscue strategies for age groups (n=104)

Age	Mean	S. D.	Mean	S. D.	Mean	S. D.
	High Graphic		High Sound		Same Gram.	
7-8	47.23	14.33	41.18	14.70	50.14	10.26
9-11	53.22	14.16	46.35	13.98	60.54**	13.43
12-14	49.71	12.20	46.14	13.75	59.14**	17.37
	Self Correct		No Self Corrects		Syn. Accept.	
7-8	14.91	11.14	68.68	14.21	63.59	16.11
9-11	24.87*	15.43	53.59*	19.99	69.79	16.76
12-14	25.00*	10.60	53.29*	10.07	77.14*	13.36
	Seman. Accept.		Meaning Change		No Mning Change	
7-8	52.45	19.79	24.55	18.99	50.50	21.98
9-11	67.42*	18.19	15.88*	17.96	61.09	22.84
12-14	65.57*	14.04	14.93*	13.02	65.64	16.81

**p<.01
*p<.05

Table 2. Means and standard deviations of miscue strategies for ability groups (N=104)

	Mean	S. D.	Mean	S. D.	Mean	S. D.
	High Graphic		High Sound		Same Gram.	
Low	50.39	13.47	44.00	13.50	58.18	14.22
Average	50.21	11.95	43.50	12.35	52.50	13.03
Above	62.44*	13.09	55.56*	14.54	65.67*	10.77
	Self Corrects		No Self Corrects		Syn. accept	
Low	21.84	13.33	57.04	16.83	67.72	16.87
Average	23.21	15.92	56.88	20.82	70.33	15.70
Above	36.11*	18.54	47.77	16.64	79.33	14.71
	Sem. Accept.		Meaning Change		No Mning change	
Low	62.88	19.61	19.39	19.03	57.19	22.30
Average	62.17	17.94	17.83	18.48	61.13	21.40
Above	78.56*	11.75	06.89	08.39	75.44	17.72
	Retelling					
Low	43.64	20.72				
Average	46.64	24.33				
Above	57.00	05.66				

$^*p<.05$

As indicated by an inspection of the standard deviations reported with the means on Tables 1 and 2 it is apparent that there is substantial heterogeneity among the students in the use of these strategies. This heterogeneity was substantiated further by the visual inspection of the students' individual profiles.

Examining the means on both tables reveals that the use of self-corrects was relatively low. However, it is important to note that the use of language and meaning was relatively high. It is most probable that self-correcting behavior was low because students did not need to go back and change their oral reading because their version of the story followed semantic and syntactic rules and made sense according to the author's story line.

DISCUSSION

The analysis of how strategic use during oral reading may change across age levels supports what reading teachers would predict intuitively. Older students used language and meaning clues more frequently than younger students. But, the analysis suggests two findings that may not be predicted. First, even the youngest students used language and meaning strategies over 50% of the time. Second, across all age levels, there occurred a similar reliance on graphic and sound clues. Both of these findings suggest variability in strategy usage even among very young readers.

There was some differential performance across the ability levels of these students. That the higher ability students used all strategies more frequently than the other students is a predictable finding. It is important to note, though, that even the students who scored the lowest on the criterion measure used language and meaning strategies over 50% of the time, a much wider use of such strategies than of graphic or phonic information. It may be that, these lower level students are not as proficient as higher level students in their use of language and meaning strategies as Goodman (1967, 1969, 1973, 1985) hypothesized, but the findings suggest that good and poor readers cannot be distinguished on these strategies alone. Further, these data revealed that no single factor (e.g., graphic or phonic use) contributed to differential performance between good and poor readers.

Regardless of how the students were assigned to groups, over-reliance on graphic-phonic information was not found. This finding seems to be particularly interesting for two reasons. First, students continued to use language and meaning strategies as much as, and sometimes more than, graphic-phonic strategies even though they were enrolled in a school curriculum that was heavily phonics-oriented. Second, students' use of language and meaning strategies should not be minimized when developing a program that could capitalize on student strengths.

The variability in performance among group members as seen in the standard deviations and individual profiles seems to underscore the heterogenity of ability within poor readers and the need for individual diagnosis. Because of the variability among the students' scores, however, caution should be applied to the generalization of these findings to other groups of poor readers. Variability in selection of strategies also relates to the issue regarding poor readers' overuse of certain strategies. The heterogeneity of strategy use appeared across all strategies which, once again, does not provide support for limited use of specific strategies. If poor readers, as a group, overrelied on certain strategies (e.g., graphic or phonic clues), the scores would be more homogeneous on these strategies. Also, the individual profiles, when grouped according to ability level, would have more commonality of performance on the use of certain strategies if there was such a preference.

EDUCATIONAL SIGNIFICANCE

Too often low ability readers have been characterized as the "have nots", those who, collectively, do not have the ability to use specific strategies. Rather than differentiating good and poor readers by their use of specific strategies, oral reading miscue research can be used to clarify the nature of variability of strategy usage of students who are thought to be quite limited in their range. With careful attention to a differential diagnosis that examines carefully student ability across several dimensions, we can be more clear in the identification of students' differential use of strategies and in appropriate program planning. If we recognize that poor readers have the ability to use language and meaning strategies, as these findings suggest, we as reading teachers can maximize the use of such strategies and build upon them within a language- and meaning-centered curriculum. Building on students' sense of language and story meaning could change the orientation of remedial reading programs that tend to be "skill and drill" with little application to the language and meaning of reading.

REFERENCES

Allington, R. L., & Strange, M. (1977). Effects of grapheme substitutions in connected text upon reading behaviors. Visible Language, 11, 285-297.

Biemiller, A. (1970). The development of the use of graphic and contextual information as children learn to read, Reading Research Quarterly, 6, 75-96.

Gilmore, J. V., & Gilmore, E. C. (1968). Gilmore oral reading test. New York: Harcourt, Brace & World, Inc.

Goodman, K. S. (1965). A linguistic study of cues and miscues in reading. Elementary English, 639-643.

Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. Journal of the Reading Specialist, 6, 126-135.

Goodman, K. S. (1969). Analysis of reading miscues: Applied psycholinguistics. Reading Research Quarterly, 5, 9-30.

Goodman, K. S. (1973). Miscues: Windows on the reading process. In K. S. Goodman (Ed.) Miscue analysis: Applications to reading instruction. Urbana, IL: NCTE.

Goodman, K. S. (1985). Unity in reading. In H. Singer & R. B. Ruddell (Eds.), Theoretical models and processes of reading. Third Edition. Newark, DE: IRA, 813-840.

Goodman, K. S., & Burke, C. L. (1971). Reading miscue inventory: Manual. New York: Macmillan Publishing Co.

Goodman, K. S., & Goodman, Y. M. (1977). Learning about psycholinguistic processes by analyzing oral reading. Harvard Educational Review, 47, 317-333.

Leu, D. J. (1982). Oral reading error analysis: A critical review of research and application. Reading Research Quarterly, 17, 420-437.

Paris, S. G., & Myers, M. (1981). Comprehension monitoring, memory, and study strategies of good and poor readers. Journal of Reading Behavior, 13, 5-22.

Singer, M. H. (Ed.). (1982). Competent reader, disabled reader: Research and application. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Smith, F. (1971). Understanding reading. New York: Holt, Rinehart & Winston, Inc.

Weber, R. M. (1970). A linguistic analysis of first grade reading errors. Reading Research Quarterly, 5, 427-451.

COLLEGE READING PROGRAMS
1958-1978

ANNE L. MALLERY
Millersville University

INTRODUCTION

College reading programs have changed significantly during the last twenty years. The impetus for change was President Johnson's "War on Poverty" (1965) which made federal funds available for compensatory education at the college level.

Since reading is a vital component of independent learning and few college instructors had the training or experience in planning developmental instruction for adults, reading specialists became key personnel in these programs.

As compensatory education became more refined, educators became sensitized to the many variables that affected success in college. The complexity of providing cost effective supportive services required curricular adaptations and innovations. It soon became apparent that only a multi-faceted approach resulted in sustained performance and graduation.

It is the purpose of the paper to review the literature and compare college reading programs prior to 1965 with those that followed.

COLLEGE READING
PRIOR TO 1965

College reading programs prior to 1965 can be characterized by their diversity in organization and orientation. In a twenty-five year review of the literature, McDonald (1957) reported a steady and widespread acceptance of remedial reading courses at the college level. This growth was encouraged by studies which suggested a relationship between college success and reading ability. However, McDonald questioned the validity of this research since virtually all studies suffered from inadequate experimental design and no evidence of attempts to control variables of curriculum, age, sex, intelligence, and motivation.

Staiger (1964) contended that no college reading program could be identified as universal, since different departments utilized different methodology. English departments emphasized vocabulary, interpretation of meaning, appreciation of style, and understanding of symbolism. Psychology approaches reinforced procedures and systematic methods to increase memory. Counseling departments attempted to find solutions to students' academic problems by discussing personal and study concerns.

Shaw (1961) identified the organizational structures of college reading improvement programs as: 1. separate, special services; 2. a part of the language arts course; and, 3. an intrinsic part of each subject. He noted that the most common pattern appeared to be the special service. Programs were offered under a variety of departments: psychology, English, education, educational psychology and personnel services. Shaw (1961) categorized programs into three orientations 1. mechanical aid oriented; 2. textbook oriented; and, 3. counseling oriented. The orientation of the programs was determined by and related to the department offering the course work.

Bliesmer (1959) attested to the increasing quantity and variety of materials which emphasized a trend toward addressing a wide scope of reading skills.

Schick (1964) agreed that programs were diverse and concluded that the most effective and enduring instruction contained elements of: detailed analysis of specific student needs as evidenced by tests and observation; selection of the most useful printed matter and mechanical teaching aids to enable students to eliminate deficiencies; continued modification of instructional techniques, materials, and devices.

COLLEGE READING
1965-1980

The advent of the "new students" meant changes for institutions of higher learning. Power (1976) warned that the inability to deal with students on a personal level would cause open door admissions policies of colleges to revert to revolving doors. If students were to graduate, changes must be made in student recruitment/ admission policies, faculty behavior, and curriculum. Power identified crucial issues as: are colleges prepared to provide remedial work? Are colleges prepared to deal with the psychological, intellectual, and emotional needs of students? Are colleges willing to hire minority staff? Are colleges willing to institute open communications so that social and political channels are available to students?

Success appeared to depend upon the faculty's acceptance of social responsibility. Power recommended: 1. comprehensive testing; 2. evaluation of traditional methods of teaching; 3. attempts to develop new curriculum; 4. re-examination of the role of the college teacher; 5. separate faculty to teach remedial work; and, 6. credit for remedial courses.

Fincher (1979) in a discussion of the impact of the "Back to the Basics" movement on colleges, pointed out some frustrations encountered by coordinators developing college compensatory programs. Who was to teach the basic skill courses? The field was highly specialized and few teachers were prepared to teach remedial course offerings.

In addition, these positions were viewed as low status roles by the academic community. The hours were long; the work was physically and emotionally draining. These factors contributed to high personnel turnover.

Instructional Philosophy

Diversity existed between programs at the different institutions, but Henderson (1976) identified two theoretical formats from which most programs were derived: the personalized approach based on cognitive field theory; the prescriptive approach based on behavioristic learning theory.

Personalized programs emphasized student goal setting, selection of materials, and verbal interaction with the teacher. In the prescriptive model, instruction was based on diagnostic findings; was pre-established; included laboratory materials, machines, and minimal teacher supervision. Henderson reported that findings were not statistically significant as to which approach produced greater student achievement. However, students using the personalized approach were more able to make decisions concerning reading improvement when provided with information concerning reading behavior.

Strang, McCullough, and Traxler (1967) argued the personalized approach was needed in college programs to help the broader spectrum of students who enrolled at institutions but whose academic skills were perceived as inadequate for success in college work.

Olsen and Swiss (1976) reviewed college reading programs at four settings: Harbor College in California, Utah Technical College at Provo, Parkland College in Illinois, and New York Community College in New York City. Examination of the models used substantiated Strang's findings that a sound reading program must stimulate both students and teachers and that the skill, interest and personality of the teacher contributed greatly to the success of the program.

Instructional Strategies

Melargano (1978) reported the results of a two-year study of intervention strategies devised to increase the retention of economically, culturally, and educationally deprived students in post-secondary schools. Objectives which increased retention were identified as: improving cognitive skills; maintaining an acceptable grade point average; developing study skills and habits; developing an awareness of majors; developing the ability to cope with institutional bureaucracy; developing the ability to cope with parental disinterest; developing a positive self perception of ability; developing a positive attitude toward the learning process; maintaining a low level of delinquent behavior; improving motivation to achieve; frequent contact with a counselor and assistance in financial aid.

This study is representative of others which pointed out the complexity of providing compensatory education at college levels. Developmental programs met this need by broadening in scope to focus not only skills, but to address the institutional, bureaucratic, and curricular influences which impede or enhance success.

An example of the expanding role of reading in the college curriculum was demonstrated by a study of the reading programs at 97 two-and four-year colleges in the state of California. Wassman (1977) studied college catalogs and made on-site visits in assimilating the following data:

1. Two-thirds of the colleges required enrollment in reading programs
2. Seventy-seven percent of the reading programs were offered through English departments
3. Fifty-nine percent of the programs provided lectures with labs
4. There was a wide diversity in scheduling mini courses
5. Forty-one colleges offered reading and study skills for students at all levels of development
6. There was a paucity of testing materials. The Nelson Denny Reading Test was used most frequently
7. Alternative programs were being offered: Saturday classes, Friday night/Saturday morning classes, Sunday classes, phone-a-course
8. Seventy-four percent of the courses were titled Developmental Reading
9. Speed reading was not a part of most programs
10. All colleges visited felt that reading and study skills were inseparable
11. Vocabulary and spelling courses were being offered in combination
12. Programs were attempting to meet community needs
13. There was little evidence of computer assisted instruction

Wassman concluded that the ideal program was eclectic, providing testing, a variety of materials, and an instructor who understood the adult learner.

Chaplin (1977) and Shepherd (1977) recommended a broad based reading program providing instruction for all college students. To help promote development of these programs, Chaplin suggested methods which can be employed by reading personnel in soliciting support from the academic community.

1. Dispel the myth of remediation which separates reading programs from the traditional academic departments. Emphasize the relationship of reading to cognitive growth; reading growth as a continum; reading as a necessity to all academic instruction; reading woven carefully into curriculum fabric

2. Define and articulate reading philosophy. Focus on what students are capable of doing rather than standardized test results

3. Make an effort to broaden the population base

4. Correlate content of the reading program to content in the academic courses

The University of Northern Colorado conducted a series of seminars and workshops for university professors and high school teachers to work cooperatively in developing instruction to improve the reading and writing of students through subject area content. In the seminars, emphasis was placed on the value of language common to all students enrolled in content courses. Academic success requires understanding of language that controls actions (directions, experiments), orders thinking (definitions, arguments) and promotes abstraction (theories, poetry).

In this approach, teachers identify the type and amount of language that fits their content area, construct informal assessments of students' abilities, and select methods/materials for clarifying reading and writing assignments. Pavlik (1976) contended that improving the reading/writing instructional practices of university and high school teachers improves the reading/writing abilities of university and high school students and future teachers.

Currie (1978) stressed that the importance of a good referral system by stating that a good referral system makes the difference between a stagnated, struggling remedial program and a flourishing comprehensive program which plays a dynamic part in the college curriculum. She emphasized that programs must be flexible. Students must be able to enter any time during the semester. The program should be individualized, personalized and content oriented.

Currie also suggested that referral systems can be strengthened by orienting faculty advisors, serving as a consultant for content area faculty, and issuing college wide publicity about reading services.

Much discussion has ensued concerning testing the reading of the adult learner. Smith (1977) reported that most college reading programs rely on survey rather than diagnostic tests to assess the reading abilities of students. In an effort to evaluate the impact of these measurements of student achievement, Smith conducted a study with 75 students. Diagnostic testing was administered to one group of students in the area of comprehension and phonics; the other group was tested with a survey test. Results after individualized instruction produced no significant differences. Smith concluded that college reading instruction cannot increase student achievement by diagnosing comprehension and phonics skills and recommended that a new method of diagnosis is needed to assess a wide range of reading, learning and study skills.

The issue of motivating students has always been a priority with teachers of adult readers. Shenkmen (1977) helped Bronx Community College students realize the relevancy of reading instruction through "Reading, So What? Workshops." Students were divided into groups and through pantomiming and gaming activities concluded that: 1. non-verbal communication is inadequate; 2. language must be written down to communicate complex ideas without distortion; and, 3. complex directions require a technical vocabulary.

Flynn (1977) motivated Winthrop College students in a Reading and Related Study Skills course under the guise of speed reading. She felt that many students do not utilize flexible rates in reading, so she taught skimming, previewing, and predicting in an attempt to dispel the guilt for not reading every word. This instruction was balanced with notetaking, study procedures, and reading skill instruction.

Content reading was becoming more prevalent in college reading programs. Monteith (1978), Mayfield (1978) conducted a study of learning centers and college reading instructors and was able to identify thirty-three content reading programs. The methods for implementing these programs varied: 1. reading/content team teaching; 2. mini courses; 3. integrated reading instruction; 4. adjunct courses; and 5. self-help labs on how to study. Content reading was also reported in the literature by Bergman (1977), Guilford (1977), Wilson (1978), and Kettlewell (1978).

Brigham Young University offered a non-credit course to law students to provide instruction in reading law cases. Students were taught skimming, scanning, phrase reading and study formulas. Mayfield (1978) increased students' comprehension by a method he called F.A.I.R. In reading a case, law students highlighted the facts, action taken, issues, and reason for the decision.

At Queensborough Community College, Bergman (1977) used a variety of instructional approaches to teach reading in content areas: synchronized instruction, team teaching, mini courses, and workshops to teach specific skills or processes for content classes.

Guilford (1977) conducted a study to determine the affects of supplementary content materials on the reading achievement of science and technology students at Texas Southern University. Reading activities were developed from journal, newspaper, and magazine articles. Pre-and post-testing revealed significant gains in the reading scores of students using the content materials.

Wilson (1978) reported the results of a reading/study skills course conducted for students who wanted to learn how to study for a history course. Goals of the course were: 1. improved study strategies using history as a medium; 2. application of these strategies to other social science courses; 3. understanding of one's self as a student; and, 4. experience in group study. After thirteen sessions, students completed a questionnaire and grades were collected to evaluate the effectiveness of the course. The mean grade of the experimental group was 3.2; the mean grade of the control group was 2.7. Results from the questionnaire ranked the value of the course highly. Based on these data. the history department at UCLA added an experimental section of the history/study strategy course to the curriculum.

Tidewater Community College provided a multi-faceted reading program designed to address the reading/learning needs of developmental, correction, remedial, adopted and accelerated readers. Kettlewell (1978) believed that since reading is a skill development process which all students learn individually, college programs must be designed to enhance learning for students at all levels of development. The Kettlewell model provided differentiated instruction for students categorized by a set of criteria. Consulting services were provided for faculty in the form of workshops in understanding the reading process, vocabulary, study procedures, and test taking strategies.

A time-line of the development of college reading programs, 1958-1978 is provided to help the reader in perceiving highlights of these two decades.

CONCLUSIONS

One of the priorities of education is to develop individuals who are independent learners. Technological and social change occurs so rapidly that information becomes outdated soon after it is taught. While few educators would disagree with this concept of teaching process skills, the difficulty has always been selecting the techniques which would achieve this purpose.

When the federal government made funds available for university compensatory education programs, reading specialists were hired to analyze the problems of the underachiever and devise alternative strategies for learning. Through the planning and evaluation process, educators examined the many variables that affected the performance of adult learners: background preparation in the subject area; balanced schedule of courses; the ability to predict and take tests; sequencing courses; note-taking; study practices; interaction between schools/colleges; and, family/faculty/program support.

The concept of college reading instruction had to be expanded to meet the demands of the task, enabling students to master the process which would help them to become independent learners. College reading instruction could not be effective without a multi-faceted approach.

As compensatory education programs developed strategies for teaching processes required for success in traditional courses, the college community became aware that instruction provided to all students could raise university retention statistics.

In addition, the political arena of college campuses favored growth of support for compensatory education. College enrollments were dropping and legislators balked at raising additional revenue for higher education. If reading programs could enable more students to graduate, colleges would be financially healthier.

Instructors of college reading analyzed the curriculum and began to work with professors, the content, the institution, the community and the students. Instruction was customized to fit the goals of the institution and the individuals. Models (e.g. Kettlewell) provided a variety of services to the universities. They refined the skills of weaker students, challenged the stronger students and provided inservice programs for faculty. The fragmented skill based programs of the past gradually evolved into university programs which offered a broad base of academic and supportive services.

REFERENCES

Bergman, I. (1977). Integrating reading skills with content in a two year college. Journal of Reading, 20, 288-291.

Bliesmer, E. P. (1959). Future practices in reading instruction in adult programs. In J. A. Figurel (ed) Reading in a Changing Society. International Reading Association Proceedings (IV), New York: Scholastic Magazines.

Carnoy, M. (1977). Schooling in a corporate society. New York: David McKay Co., Inc.

Chaplin, M. T. (1977). Where do we go from here? Strategies for survival of college reading programs. (ERIC Document Reproduction Service No. ED 147 791)

Currie, S. R. (1978). Referring community college students to the reading program. Journal of Reading, 21, 409-415.

Fincher, C. (1979). What are the basics and why the concern? Athens, Georgia: Georgia University. (ERIC Document Reproduction Service No. ED 171 224)

Flynn, P. Speed is the carrot. Journal of Reading, 1977, 20, 683-687.

Guilford, M. C. (1977). Using content area periodicals to supplement college reading instruction. Reading Improvement, 14, 172-174.

Henderson, M. A. (1965). Individualized reading instruction in two-year colleges. Journal of Reading, 19, 464.

Kettlewell, G. B. (1978) All college students benefit from reading instruction. (ERIC Document Reproduction Service No. ED 159 622)

McDonald, A. S. (1957). Influence of a college reading improvement program on academic performance. Journal of Educational Psychology, 48, 171-181.

Mayfield, C. (1978). Reading in content areas in higher education. Unpublished report. Brigham Young University, Provo, Utah.

Melargano, R. J. 1978). Planned variation study: Intervention strategies for secondary compensatory education. Santa Monica, California. (ERIC Document Reproduction Service No. ED 154 677).

Monteith, M. (1978). Beyond basic skills courses in colleges to courses in basic concepts and content area reading. Journal of Reading, 22, 74-77.

Olsen, T. & Swiss, T. (1976). College reading programs. Journal of Reading, 19, 523.

Pavlik, R. (1976). Improving reading and writing in the content fields; High schools and colleges can cooperate. Journal of Reading, 20, 237-240.

Power, M. (1976). New students mean change for colleges. Journal of Reading, 20, 237-240.

Schick, G. B. (1964). Diversity in college reading programs. In P. Leedy (Ed.), College Reading Instruction. Newark: International Reading Association.

Shaw, P. (1961). Reading in college. In Sixtieth Yearbook of the National Society for the Study of Education (Part I), Chicago: University of Chicago Press.

Shenkman, H. (1977). Reading, so what? Journal of Reading, 20, 567-569.

Shepherd, J. F. (1977). The failure of non-credit college reading courses. Journal of Reading, 20, 493-497.

Smith, B. D. (1977). Do we need differential diagnosis at the college level? No. Journal of Reading, 21, 62-66.

Staiger, R. C. (1964). Humanistic aspects of college and adult reading. In P. Leedy (Ed.) College-Adult Reading Instruction. Newark: International Reading Association.

Strang, R., McCullough, C. M. & Traxler, A. E. (1967). The Improvement of Reading. New York: McGraw-Hill Book Co.

Sullivan, R. J. (1975). Learning disabilities in college: The university role. Journal of Reading, 19, 6-7.

Wassman, R. (1977). California college reading programs: The state of the state. National Institute of Education, HEW, 1977. (ERIC Document Reproduction No. ED 146 559).

Wilson, M. M. (1978). A report on an adjunct course designed to integrate reading and study skills with a university academic class. (ERIC Document Reproduction Service No. ED 159 611).

Time-line for the Development of College Reading Programs 1958-1978

1957	McDonald pointed out flaws in the design of studies of college reading programs.
1959	Bliesmer reported an increase in the quantity and variety of materials which indicated a trend towards addressing a wide scope of reading skills.
1961	Shaw grouped programs by three orientations: mechanical aid oriented, text book oriented, counseling oriented.
1964	Schick agreed that college reading programs were diverse and identified the characteristics of effective and enduring programs.
1964	Staiger contended that no college reading program was universal since different academic departments focused on different methodology.
1967	Strang et. al. favored the personalized approach and linked success to student aspiration and teacher attitudinal factors.
1975	Sullivan recommended the personalized approach to address the needs of a broad spectrum of students.
1976	Henderson identified two theoretical formats from which programs were derived: the personalized and the prescriptive.
1976	Power reinforced Sullivan's preference for the personalized approach rather than the behavioristic model.
1976	Olsen and Swiss substantiated Strang's findings that a successful reading program must stimulate both students and teachers.
1976	Pavlik contended that improvement in the reading and writing instructional practices of teachers would improve the reading and writing abilities of the students.
1977	Carnoy linked student success to the skill, interest and personality of the teacher.

1977	Wassman concluded that the ideal reading program was eclectic, providing testing, a variety of materials, and an instructor who understood the adult learner.
1977	Shepherd suggested increasing the credability of reading instruction as a vital part of the college curriculum.
1977	Chaplin emphasized the continuing importance of reading growth in cognitive development and academic instruction.
1977	Smith recommended that a new method of diagnosis is needed to assess a wide range of reading, learning, and study skills.
1977	Shenkman presented a program for students illustrating the importance of written language and an extensive vocabulary to communicate complex ideas.
1977	Flynn described a reading course which incorporates notetaking, study procedures and reading skill instruction with speed reading such as skimming, previewing and predicting.
1977	Bergman favored a variety of instructional approaches to teach reading in the content areas.
1977	Guilford determined that the use of supplementary content materials resulted in significant gains in reading achievement.
1978	Melargano reported factors which draw attention to the importance of a strong academic/reading component to the student's overall achievement.
1978	Currie stressed the importance of a good referral system to the success of a reading program and recommended faculty advisor orientation, distribution of publicity about reading services, and reading personnel serving as consultants for content area faculty.
1978	Monteith reported that content reading instruction is becoming prevalent in college reading programs.
1978	Mayfield reported content reading program for law students.

1978 — Wilson reinforced the positive effect of content reading skills programs on grades in that content area.

1978 — Kettlewell provided a model for an individualized program for reading skill development with consulting services provided for the faculty.

1979 — Fincher pointed out frustrations which are encountered by coordinators developing college compensatory programs.

DEVELOPMENTAL READING PROGRAM INNOVATIONS AND PRACTICES

TERRY BULLOCK AND DEBORAH ROSECRANS MADDEN

University of Cincinnati

The Reading and Study Skills Program at University College, University of Cincinnati has served students and faculty for twenty years. During this time, the program has gone through a number of changes due to funding, student needs, and administrative priorities. For the last five years, the program has offered a number of innovative courses and services for students in University College, as well as for students from other colleges and students not yet in college. This paper provides a descriptive overview of the Reading and Study Program's innovations and practices. The innovations include these areas:

1. Placement Testing and Advising
2. Instruction Practices
    - A. Paired Reading-Psychology Courses
    - B. Pre-Technical Program Developmental Courses
    - C. College of Education Developmental Courses
    - D. Computer Instruction
3. Future Innovations
    - A. Clinic-Entering Freshmen

Summer Placement Testing and Advisory Procedures

All incoming University College freshmen take reading, writing, and math tests to determine their placement in developmental courses. The DRP (Degrees of Reading Power) Test - Form PA2 - is given during the ten-week summer period to determine if students should be placed in Effective Reading fall courses.

The DRP Test, developed by the New York State Education Department, Touchstone Applied Science Associates, and College Board assesses a student's ability to comprehend prose at increased levels of difficulty. The DRP Test is a modified cloze test where seven words are deleted from each 325-word prose passage in the test. The student has to select the correct word for each deletion from five possible alternatives.

The Reading and Study Program uses a 64 unit score (71 raw score) to screen students because 64 is equivalent to reading material at the 12th grade level according to the College Board Readability Report (1981). Four units, (example - 64 to 68) indicate approximately a one grade change. The student's reading level and readability level of reading material can be equated using the unit score designation. In forty-five minutes, a student must correctly answer 71 out of 98 questions to be evaluated as having adequate reading skills for college work.

A positive feature of the DRP Test is that it tests reading holistically, measuring the skills of comprehension, vocabulary, and rate of reading. Also, the test results can be easily explained, testing time is short, administration is easy to large groups, and sequential test item difficulty is easier to explain to students than in tests with skills divided into different test sections. This test is one of three given during the Summer Placement and Advising Program. Students must also take a twenty minute writing test followed by a forty minute math test. In the afternoon session the students meet individually with a faculty advisor to determine their fall courses and major program requirements.

Paired Reading and Psychology Courses

The innovative practice of pairing a reading class with a content class has been done at the University of Cincinnati for the past five years. The pairing concept involves the content oriented class Introduction to Psychology and the "how to" Effective Reading and Study Skills course. These courses together form a 5 credit hour class where the communication and study strategies taught in Effective Reading are biased towards Psychology. The impetus for pairing came from the psychology instructor who felt that most of the students in his classes receiving D's and F's could benefit by having a reading and study course paired with the psychology class. Additionally, the psychology instructor did not feel that he could spend that much class time teaching notetaking, test taking, and other reading and study strategies.

There are several factors that make this an innovative paired offering. First, the reading instructor attends all the psychology lectures. This enables him to assist students more directly. For example, he was able to model the Cornell Notetaking System by using the actual notes from the Psychology class. It also enables him to assist students with such things as anticipating test questions and reviewing for the Psychology quizzes. Second, a number of specific materials have been developed and tried out over the past several years of this course including study guides covering the material in the psychology textbook. Also, many of the activities that have been used with students have been put into the Instructor's Manual - Study for Success. Activities such as concept trees, time lines, schedules, and anticipatory test questions were all implemented and refined in the Reading class.

Besides the innovations described above, reading test data and retention data for the program were collected (refer to Fifth American Reading Forum Yearbook, 1984, p. 89). The data gives the clear indication that the paired course really does make a difference both in terms of Psychology test scores and retention from freshman to sophomore years.

Additionally, during the 1984-85 year, the Human Psychology classes were paired with reading courses as follows: Human Motivation was paired with two reading sections in the fall and Human Development was paired with one reading section in the winter. These are the first two courses in the three-part Human Psychology series which also includes Human Relations. The three Psychology courses are a requirement for approximately half of the University College population. For example, the Pre-Nursing students in University College must take all three of the Human Series to transfer to the College of Nursing and Health.

Differences in the three Human series and Psychology I and II paired are these:

1) The three Human Psychology classes do not carry the same course number for the content and reading courses. Students in the Human series voluntarily enroll for the paired reading courses. During Fall Quarter, 1984, twenty students took the paired reading course with Human Motivation (ten students in each section). In the Winter Quarter Human Development course, sixteen students were enrolled in one section. The additional number of students (from ten to sixteen in one session) indicates students' sense of need for the paired reading course.

2) The psychology instructor's evaluation method was the same in the three classes, making the paired courses' objectives easier. Also, grades increased for the paired Human Development students who had also taken

paired Human Motivation. Clearly, use of the reading and study skills in both classes had been one factor in helping students to maintain and increase their ability to learn content materials.

Pre-Technical Program

The Pre-Technical Training Program is a cooperative effort between the University College - University of Cincinnati, the Private Industry Council (PIC) of Cincinnati and Hamilton County, and the Employment and Training Division of the City of Cincinnati. The program provides academic preparatory education opportunities for Job Training Partnership Act (JTPA) eligible adults and out of school youth who are unable to meet the requirements for admission into a two-year technical associate degree program because of their need for improvement in communication skills.

During the program's duration of two or three quarters, students' curriculum consists of courses in developmental math, reading, English, and other preparatory courses in Career Planning, Consumer Economics, Practical Logic, and Introduction to Computers.

The program then, is designed for non-traditional aged students who are returning to school. The students go through a transition year where they pick up the necessary skills and prerequisites to enter into either a technical or academic program. The focus of the Pre-Technical reading and study class is on developing these skills that will carry over into content area classes. For example, the skills of managing time and breaking large tasks into small manageable ones receives a great deal of emphasis. The students benefit from developing a weekly schedule and prioritizing tasks. Additionally, most of these students have always tackled a textbook chapter by reading the entire thing rather than breaking it into manageable segments as is shown in the SQ3R (Survey, Question, Read, Recite, and Review) Method. Besides these skills areas, a great deal of time is spent on reading a variety of materials and discussing them in class. The students benefit from these exchanges by gaining self-confidence when contributing their ideas in a class discussion situation. Finally, vocabulary development has received a great deal of course emphasis since these students need to be able to handle the demands of various content materials. Specifically, vocabulary work in this area centers on context clues and prefixes.

College of Education Course

Another special offering has been a course designed for education majors who have not passed the reading portion of the Pre-Professional Skills Test. The reading staff has designed a course for these students which concentrates on vocabulary and comprehension development. In particular, the staff works with these students on reading passages and answering main idea and inferential questions. The majority of students have the greatest difficulty in answering inferential questions. In the area of vocabulary, most of the emphasis is placed on context clues. The students need to quickly identify and define words that they encounter in test passages. The materials used in this course consist of comprehension modules developed in the program on main idea and inference. Additionally, these students take a computer assisted instruction course on main idea. Other materials used in this class consist of passages taken from a variety of content areas.

Computer Instruction

The instructional aid that can be obtained from computer usage was recognized two years ago in the Reading Program. Specifically, its individual tutoring ability was felt to best reinforce instruction. The Vocabulary Skills Laboratory course was then written and programmed on the computers in the Computer Assisted Instructional Laboratory in the University of Cincinnati. This course consists of a series of lessons in phonics, structural analysis, context clues, and study skills. Students can register concurrently for Effective Reading I and Vocabulary Skills Laboratory I in the Fall Quarter, and for Effective Reading II and Vocabulary Skills II in the Winter Quarter. The Effective Reading course is two credits each quarter, and the Vocabulary Skills Laboratory is one credit each quarter. The Vocabulary Skills Laboratory's course content consists of areas that the instructors agree can provide reinforcement for the low reading level student. For example, vocabulary by context clues and structural analysis is reviewed in Effective Reading I because students have used in the past only dictionary skills to learn vocabulary. The Vocabulary Skills Laboratory computer module provides the additional instructional assistance and reinforcement with an individual tutor for each student - the computer.

Future Directions - Reading Clinic

Beginning in the summer of 1986 it is anticipated that a summer reading program for high school age students will be offered. In the past couple of years, the reading staff has received calls from parents requesting help for their sons or daughters who have reading problems. This summer program is designed to work with 25 high school students on a 1 to 5 ratio of teacher to student. Besides the low student to teacher ratio the class will provide a range of activities and materials for each student including computer assisted instruction, specific skill development, high interest reading materials, study skills, library orientation, and reading/writing activities. In addition the program will test each student to determine specific areas of strengths and weaknesses. In this manner, the clinic will be able to more closely tailor instruction to student needs. The long range goal for this program will be to help these high school students gain the self-confidence and skills necessary for entering college.

Summary

The Reading and Study Program at the University of Cincinnati has developed a number of innovative courses over the past several years in order to provide services to various audiences. For example, the paired reading courses have been aimed at students who need specific help with their psychology classes while the developmental reading course for education majors helped students to pass the Pre-Professional Skills Test and to gain self-confidence in their communication skills. Additionally, the University has served non-traditional aged students in the Pre-Technical Program who need assistance in how to read and study for college level courses. These innovations have been bolstered by the development of computer assisted instruction software in the areas of vocabulary and comprehension. Finally, the reading staff is looking to the future with the development of a Summer Reading Clinic for high school students.

GUIDELINES AND PRACTICES OF EVALUATION OF STATE READING PROGRAMS:

HOW VIABLE ARE THE CRITERIA?

PATRICIA K. SMITH
West Virginia University

SONDRA L. REBOTTINI
Ohio University

Historically, each state has been responsible for the organization, administration, and implementation of public school instruction. Every state has a designated agency, usually a state department of education, which provides school districts with minimal competencies and/or basic skills to be taught within local schools. School districts have the freedom and flexibility to design their curricula while fulfilling the basic skills standards established at the state level. However, because of a decrease in pupil performance, evaluation of these curricula has become a major concern. Evaluation within public education has traditionally been a troublesome area (Thomas, 1969).

In the 1970's educational quality and mastery of a minimal number of specific competencies brought about a need for student performance evaluation. Since 1974 all fifty states have experienced some legislation or state board activity related to establishing educational standards (Lovell, Riegle & Bunke, 1981). By 1978, 20 states had or were considering legislation related to minimal competency testing (Chall 1979). Within this period an interest in assessing reading competencies began to develop as a means of improving literacy.

In an effort to ascertain whether state departments were becoming directly involved with the evaluation of students' reading proficiency, two university faculty members, one from Ohio and one from West Virginia, developed a written interview which was distributed to state superintendents of schools in all fifty United States. The focus of this study was to determine the type and extent of involvement that state departments of education provided to individual school districts. In addition to reviewing guidelines designated by each state, this study examined the evaluation practices of state reading programs as well as the criteria which were used for program evaluations.

Statement of Purpose

The primary purpose of this study was to compile evaluation data on a national level by contacting state superintendents, requesting information related to state reading programs. The researchers were attempting to determine the existence and nature of guidelines and criteria for evaluating reading programs. An effort was made to establish related components such as personnel, materials, and similar items which might be involved during an evaluation. Respondents were requested to indicate conditions under which state evaluations were initiated and administered; and then identify the personnel involved. The prioritizing of the identified needs and how the assessment results were used were also investigated.

The secondary purpose concentrated on the reported criteria and considered the ramifications which could result from reading program evaluations. Attention was directed toward the establishment of what had been done in the past, is currently in effect, and is proposed for the future.

METHOD

The fifty state superintendents of schools were contacted by mail and asked to respond to a written interview. An introductory letter, which was included with the written interview, explained the purpose of the study. The interview questions concentrated on guidelines, practices, and criteria used in the assessment of reading programs within each state.

Replies were received from thirty-three of the fifty states (66% return). The replies yielded a sample which represented all geographic areas of the United States. In many instances reading was included with overall achievement evaluation and not isolated as a separate discipline for evaluation. Consequently, state superintendents or their designated representatives provided information on all achievement testing. Answers were given to the interview questions and, frequently, state superintendents enclosed materials which had been developed at the state level.

Individual states tended to define "state evaluation" according to their own perceptions and did not conform to a particular standard. Some states interpreted statewide evaluation to include tests developed within the state based on minimum standards established by the state. These tests were distributed to local school districts where they were administered by local personnel. Other states defined state evaluation to mean the control of evaluation was initiated at the state level, which determined commercially prepared tests which were to be administered at the local levels.

Some states reported that guidelines and criteria were being formulated at the state level; therefore, background information concerning the development of assessment practices for state reading program evaluations was included. If no state evaluation was in effect, the interviewers were directed to contact local school districts or were given materials such as state listings of basic skills or competencies.

Of the thirty-three responding states, twenty-three (70%) reported that evaluation of student performance was originated under the aegis of the state department of education and coordinated by a division which reported directly to the state department. Seven of the responding states (21%) indicated that evaluation of student performance was the direct responsibility of local school districts who used assessment instruments of their own choosing, receiving no or few restrictive guidelines from the state. One superintendent responded that his state currently has no pupil evaluation evolving from the state level, but legislation for student evaluation was under consideration. Three states (9%) reported they did not have state evaluation of reading programs and made no mention regarding assessment of pupil performance at local levels.

RESULTS

Thirty-three of the written interviews, representing all geographic areas of the United States, were completed and returned. State superintendents or their representatives were asked to respond to the following questions of the interview: (1) What components of a reading program are evaluated? (2) What is the frequency of evaluation? (3) How is the evaluation instituted? (4) Who conducts the evaluation? (5) Who synthesizes and disseminates the information gleaned from the evaluation? (6) What is done with the results of the evaluation? (7) What types of follow-up are utilized as a result of an evaluation? For the purpose of this study, the replies were divided into three classifications: (1) respondents who reported state level evaluation of student performance, (2) respondents who indicated that the evaluation was initiated and administered at local levels, and (3) respondents who reported no state evaluation and did not mention or comment on local school evaluations.

Respondents Reporting State Level Evaluation of Student Performance.

The results of the first classification were based on responses from twenty-three states (70% of total return) who reported that some degree of evaluation of personnel, curriculum and materials was initiated and executed at the state level. More specifically, state evaluations focused on personnel with the emphasis concentrating on student rather than teacher evaluations.

What components of a reading program are evaluated? Students evaluations centered on several curriculum areas with the reading component being the one evaluated most frequently. All twenty-three states reported student evaluation of reading. Seven states (30%) indicated they assessed language arts, language skills or language mechanics but did not mention specific types of skills. However, these seven states listed reading as a separate curriculum area and did not mention whether reading was taught or evaluated within a language arts block. Fourteen respondents (61%) indicated that students were evaluated in the area of written language with the term "composition" being sepcifically mentioned by three states. Two states (9%) reported that spelling was evaluated; with one state specifically mentioned literature. Thirteen of the states (57%) volunteered the information that mathematic skills were also evaluated on a statewide basis.

What is the frequency of evaluation? The frequency of evaluation, grades tested, and method of selection of students to be evaluated were included within this question. Data, which related to frequency, suggested that the majority of states who conducted statewide evaluations were utilizing an annual testing schedule. Twenty-one of the responding states (91%) indicated that portions of their student populations were evaluated annually. One state indicated that a selected or sampled portion of its student population was evaluated on a semi-annual basis. Another state reported that, while one segment of its schools' population was evaluated on an annual basis, another segment was evaluated biennially. Two additional states reported biennial evaluation. One state reported having an evaluative measurement every fourth year.

Frequency related to grade level suggested that most states were conducting some type of evaluation across the age/grade spectrum of K-12 grades. Two states did not provide information concerning grade levels which were tested. The grade level most frequently evaluated was grade eight with fourteen states (61%) reporting assessments. Third, fifth and eleventh grades were the next most frequently assessed with thirteen states (57%) reporting. Continuing, in descending order, the following grade levels were evaluated: sixth and twelfth, 11 states (48%); fourth and ninth, 10 states (43%); tenth, 8 states (35%); seventh, 6 states (26%); second, 5 states (22%); first grade, 3 states (13%); and kindergarten, 1 state (4%). Figure 1 illustrates the frequency of evaluation from kindergarten through twelfth grade.

The final factor within this question was the method of student selection. Of the twenty-three states who reported statewide testing, sixteen (70%) evaluated students in each school system, but individual grade levels which were tested were determined by the state. Three states (13%) randomly selected school systems annually within the state. Of these three, one used a stratified random selection to represent the type of school system as well as the population areas (urban, suburban, rural). One state relied on voluntary participation of school systems. Three states did not reply to this item.

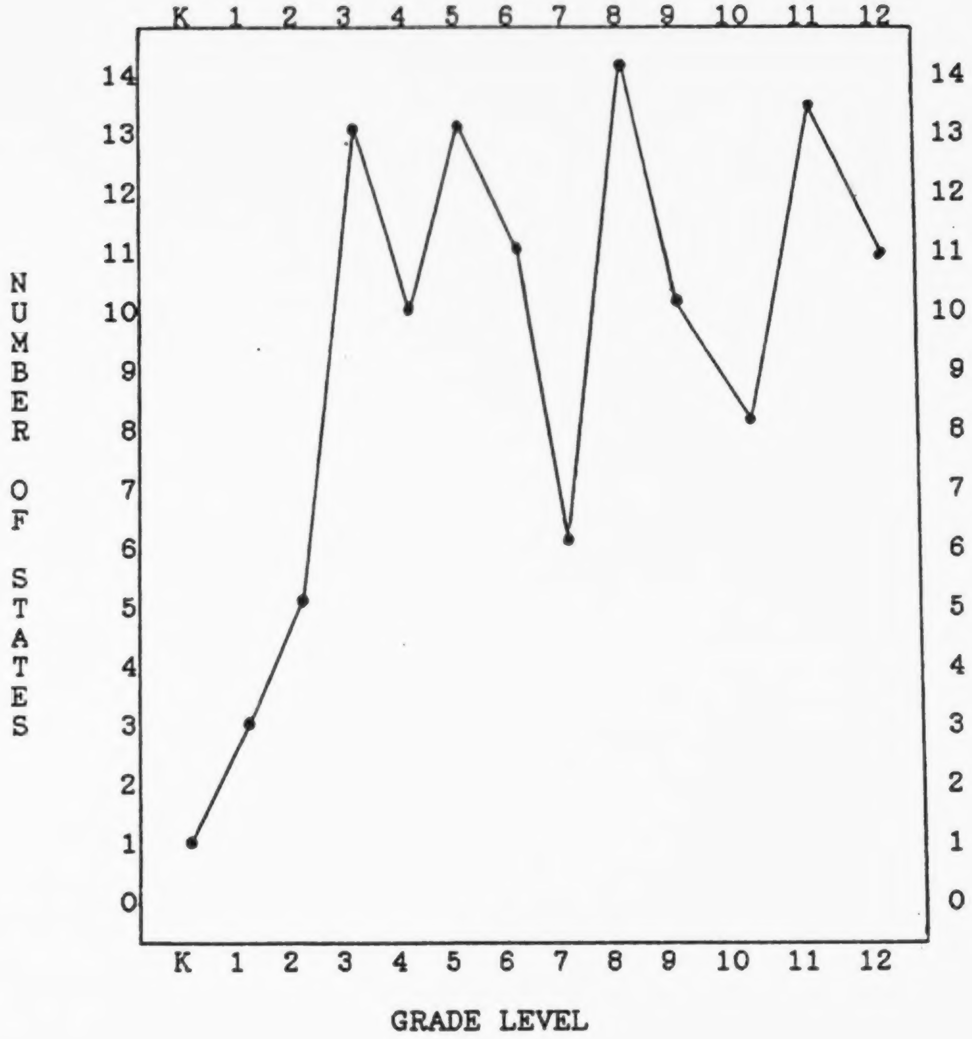


Figure 1. Frequency of evaluation by grade level.

How is the evaluation instituted? The process involved for the establishment of state evaluation was similar in each of the twenty-three states. In general, the structure of the evaluation programs was designed by a division or agency within each state department of education. Evaluation materials, procedures, timelines, guidelines, criteria levels, and at times training for the teachers or counselors, were under the auspices of a division of the state department of education.

In certain instances the evaluation was initiated at the state level with input from local levels. In other situations local school districts decided whether or not they wished to have their students evaluated and then contacted the state department of education. One state reported that annual evaluation was contingent upon funding.

Who conducts the evaluation? Although evaluations were initiated and established at state levels with the student population being identified by various methods, local school systems conducted and administered the evaluation. Despite the fact that evaluations were controlled by the state, most students were evaluated at local levels under the guidance of local teachers and administrators. Two of the twenty-three states (9%) indicated that training was provided for those individuals at local levels who administered the tests. Scoring was done by independent out-of-state boards, machine-scored by test publishers, central agencies, or the state department.

Who synthesizes and disseminates the information gleaned from the evaluation? When asked to respond to this question, six states (26%) replied that the evaluation information was synthesized and disseminated by their state departments of education. Twelve replies (52%) indicated that a committee under the supervision of the education department synthesized the results which were distributed by the state department of education. Two states (9%) reported synthesis and dissemination by the administrators of local school districts. One state reported that the evaluation information was synthesized by the state with the results being disseminated by a combination of school districts, while a second state reported synthesis by the local school district and dissemination by the state department of education. One state did not respond to this item.

What is done with the results of the evaluation? All twenty-three states having state evaluations reported that the results were used at state and local levels. Some of the states were more specific concerning the distribution of results, and indicated that the results were shared with news media, reading specialists, teachers, parents, students, and other groups who might benefit from the information provided.

Table 1

Synthesis and Dissemination of Information from Evaluations

Number and Percentages of States Reporting		Synthesis	Dissemination
6	(26)	State Department	State Department
12	(52)	State Dept. Committee	State Department
2	(9)	Local School Districts	Local School Districts
1	(4)	State Department	Combination of Sch. Dists.
1	(4)	Local School Districts	Local School Districts
1	(no response)		

Results were given to educational decision makers who were involved in state and local needs regarding program improvement. The outcomes of the evaluations determined curriculum adjustments and the establishment of goals for state and local levels. One state provided team and/or individual visits to assist the teachers and local schools in improving their reading instruction. Another state used the results for student promotions to the next grade level and also as a determinant for high school graduation. In most instances the results indicated an individual's as well as group performances. Results were frequently compared with national evaluations.

What types of follow-up are utilized as a result of an evaluation? The procedures which followed state evaluations were many and varied. The methods of follow-up fell into two classifications: those controlled by a state department committee and those controlled at the district or local levels. The type of follow-up at the state levels were varied and included: (1) test revision to conform to what is being taught in local schools, (2) establishment of levels of minimal competencies, (3) the development of special publications reporting the results of evaluations, (4) certification of schools meeting the state's minimal standards, (5) identification of state trends in educational achievement, and (6) the improvement of educational opportunities for each child. At the district or local levels, the follow-up was also diverse and encompassed: (1) the assessment of student needs, (2) identification of problems in order to provide remediation assistance and programs, (3) program evaluations to assist in the revision of course content, methods of teaching, and the schools' curricula, (4) the planning of technical assistance, (5) maintenance of staff and staff development programs, and (6) plans for future activities to improve performance.

Respondents Reporting Local Level Evaluation of Student Performance. The second classification for this study involved those respondents who indicated that student evaluation was initiated and administered by local schools. Seven of the thirty-three responding states (21%) reported that the evaluation of student performance was the direct responsibility of local school districts. One of these states indicated that local school districts determined the curriculum area, frequency, and grades to be evaluated in addition to administering, synthesizing, and disseminating the information. The results and types of follow-up were also controlled by the local districts. Another state indicated that a program assessment instrument was distributed by the state department of education but indicated that there were no state standards for assessing student performance. A third state reported that a state level English Task Force was formed to develop guidelines and competencies but the local schools operated autonomously and were not bound to the guidelines.

The state board rules for textbook adoption was submitted by a fourth state which indicated that the evaluation and selection of texts were based on a state list of required skills for graduation and promotion. A language curriculum guide was sent by a fifth state but no formal state evaluation of student performance was incorporated. Once again local school districts assessed student performance. The sixth and seventh states reported no state level evaluation but submitted booklets regarding standard skills and basic competencies determined at the state level.

Respondents Reporting No State Evaluation, No Comments on Local Evaluation. Three states responded to the cover letter which accompanied the written interview but did not answer the specific questions which dealt with state evaluation criteria and procedures. One state reported there was no state evaluation of student performance nor was any information provided concerning local evaluations. One state indicated that high schools were accredited through Northwest Accreditation which evaluates programs at the secondary level but made no mention of elementary students' evaluations. Another state had proposed legislation for grades three, seven, and ten but preferred not to answer the questions until the legislation was approved. However, the proposed criteria for student assessment was provided in an accompanying curriculum guide.

DISCUSSION

One of the purposes of this study was to determine the existence and nature of guidelines and criteria for evaluating reading programs. The returned interviews, representing all geographic areas of the United States, indicated that a majority of states have identified similar specific basic skills which are to be taught within classrooms. However, not all states have developed statewide assessments or adopted commercially prepared tests to analyze students strengths and weaknesses.

Of the twenty-three states who reported statewide testing, only ten (43%) provided background information concerning how their state's evaluation came into being. The development of basic skills within each of these states came as a result of state legislation. Subsequently, the legislators mandated testing programs to measure the skills taught. The tests to be used at specified grade levels were either prepared at the state level or commercially.

A second purpose considered the ramifications which could result from a statewide reading program evaluation. A predominant issue was the establishment of statewide standards for reading achievement and the resultant influence on teachers' and schools' efforts. Evaluation of reading as mandated by legislation is designed to provide accountability to students, parents, schools, and the general public. Most importantly, a statewide evaluation is intended to assure that students master the skills they need to become productive members of our society. Furthermore, educators are able to identify specific reading problems which enables schools to provide remediation.

Approximately 31% of the states reported legislation enacted to establish statewide evaluation. An interesting note involved the recent dates of legislation; seven states reported legislation passed during the 1970's; three states had legislation between 1982-1985; reflecting a recent surge in states' involvement in reading program evaluation. Only one state indicated that legislation is pending.

CONCLUSIONS AND SUMMARY

Approximately two-thirds of the state departments of education replied to the written questionnaire distributed to each of the United States. Of these replies twenty-three reported their state education department initiated and controlled student evaluations in their states. The conclusions and summary which follow were based on the responses which came from states representing all geographic areas.

--The states which conducted statewide evaluations assessed reading as well as other content related subjects.

--Of the twenty-three states conducting statewide evaluations, reading skills comprised the component most consistently measured. Student performance related to written communication and mathematic skills were also highly valued as components which were assessed.

--The majority of responding states reported that statewide evaluations were conducted annually. However, while grade levels to be evaluated within each state were specifically designated, populations were randomly or voluntarily selected.

--In statewide evaluations, materials, procedures, timelines, and criteria guidelines were instituted at the state level but administration of the assessments was conducted within local school systems.

--The methods of synthesis and dissemination of evaluation results were many and varied with little commonality.

--Results of state evaluations were used at state and local levels to establish reading goals, improve reading programs, and/or as a determinant for promotion and high school graduation.

--The types of follow-up encompassed long and short range goals; the predominate goals included state test revision, establishment of minimal competencies or standards, school certification, identification and remediation of students' needs, revision of course content, teaching methods and school curricula.

This study was designed to determine the effects of state reading programs. The nature of the responses has led the researchers to conclude that states are currently in the process of determining criteria for evaluation of their state reading programs. The viability of their evaluations cannot be assessed until all fifty states report reading program evaluations in a format which can be consistently applied.

REFERENCES

Chall, J. S. (1979). Minimum Competency in Reading. An Informal Survey of the States. Phi Delta Kappan, 60, 351-352.

Gambrell, L. B. (1985). Minimum competency listing and programs in reading: A survey of the United States, Journal of Reading, 28, 735-738.

Lovell, N. B., Liegle, R. P., Bunke, C. R. (1981). Minimum Competency Testing: Hopes, Fears, and Fallacies, The Educational Forum, 45, 199-206.

Stedman, C. (1985). Testing for Competence, The Educational Forum, 49, 199-210.

Thomas, M. D. (1969). Performance Evaluation of Educational Personnel, Fastback 135, Bloomington, Indiana: Phi Delta Kappa Educational Foundation.

TEACHER PERCEPTIONS OF MICROCOMPUTER COURSES

PEGGY E. RANSOM
Ball State University

In the spring of 1982, the elementary education department at Ball State University formed a committee to draft a set of microcomputer courses for teachers. This group decided to develop five courses, in conjunction with the computer science department, designed to provide students with a Ball State microcomputer teaching endorsement upon completion. The courses adapted and accepted by the University Senate included the following:

Elementary Education	Microcomputer Workshop	2 quarter hours
Computer Science	BASIC Programming	4 quarter hours
Computer Science	BASIC Programming	4 quarter hours
Elementary Education	Microcomputers in the Classroom	4 quarter hours
Elementary Education	Practicum	4 quarter hours
	TOTAL	18 quarter hours

The first course, Microcomputer Workshop, was composed of lectures, hands-on experiences, and tests. This course was designed for teachers to learn how to do simple programming, work with software, and be knowledgeable about the components of a microcomputer and its use in the classroom. Since it was a 2-hour course, it was hoped its format and time would appeal to the teachers (elementary, secondary, college, and administrators), as well as stimulate them to use microcomputers.

The purposes of the first study being reported were to determine (a) if a 2-hour microcomputer course would gain positive acceptance from graduate students, and (b) if teachers' attitudes could be changed concerning use of microcomputers in the classroom.

The author taught the beginning microcomputer course winter and spring quarters of 1983-84. A formulative/narrative survey was prepared and administered to students in the classes at the end of each quarter. A total of 95 (elementary, secondary, and administrators) were enrolled, and all of them responded to the survey. The items and results on the survey (showing percent responding in each category) are found in Chart 1.

From the first survey it was found that the teachers did enjoy the course (99%), and would recommend it to other teachers (94%). They liked the lesson "hands-on" work (94%) and did develop some ideas of how to use the computer

Chart 1

Teacher Perceptions Regarding One Computer Class

Reflecting on the microcomputer course you had at Ball State University, how would you answer the following:

1. Are the objectives of this course as stated in realistic terms?

Very much				Not at all
78%	16%	6%		

Comments:

2. As a teacher, rate your knowledge (when you began the course) of microcomputers.

Knowledge complete				No knowledge
	7%	8%	21%	64%

Comments:

3. Did you have programming skills with microcomputers before this class?

Very good skills				None at all
		3%	13%	80%

Comments:

4. Did you know how to lead programs into the computer with cassettes or discs before this class?

Could do with ease		Might be able to		Not at all
14%	7%	10%	16%	53%

Comments:

5. Did you enroll in this course with a bit of apprehension?

Completely				Not at all
32%	20%	26%	14%	8%

Comments:

6. Did you develop some ideas of using the computer in the classroom as a result of this class?

Very much				Not at all
53%	27%	18%	2%	

Comments:

7. Would you recommend this course to other elementary teachers?

Completely				Not at all
72%	22%	5%	1%	

Comments:

8. Have the "hands-on" practice work of this course been suitable for elementary teachers?

Completely				Not at all
73%	22%	5%	1%	

Comments:

9. My general attitude toward this class is:

Positive				Negative
72%	27%	1%		

Comments:

Narratives:

10. What is there about this class that is least appealing and rewarding to you?

Responses: Testing (35); not enough time (10); programming (3); filmstrips (1)

11. What about this class is most appealing and rewarding to you?

Responses: Learning BASIC programming (20); "hands-on" format work (34); enthusiastic patient professor (23); course content (3); gaining confidence (3); ideal pace (6); no pressure (1); learning commands and what they mean (1)

12. What changes would you suggest?

Responses: More class sessions (8); more software (6); practice with discs (5); no tests (3); have a textbook (3); homework assignments (2); more programming (4)

in the classroom (80%). Most came to the course with little or no microcomputer background (93%), while about half were apprehensive about taking the course (52%).

The two purposes set forth in this study were answered: (a) This microcomputer workshop was accepted positively by graduate students, and (b) positive attitudes toward microcomputers was evidenced. In addition, the course appeared to appeal to a variety of people since there were elementary and secondary teachers as well as administrators enrolled. Overall, the survey results were encouraging concerning the perception of teachers toward the beginning elementary education microcomputer course at Ball State University.

Teachers' Perceptions Regarding Five Microcomputer Courses

Second Study

By the autumn of 1985, 32 students had finished the 18 hours of microcomputer courses (listed as requirements for Ball State microcomputer teaching endorsement). These students were all elementary/middle school teachers. While teaching the last course (Practicum) to these people, several questions emerged that called for answers. How did they feel about the microcomputer courses now? How did they feel about the three elementary courses which made up 10 quarter hours, and the two computer science courses that involved 8 quarter hours? Were they using the microcomputer programs they devised in these courses in the classroom? Would they recommend courses to other teachers? Thus, the purposes of this second study were to investigate the attitudes, concern, and opinions of the students toward microcomputer classes offered.

For this study, another formulative/narrative survey was designed to meet the purposes set forth. It was mailed to the 32 students after the spring quarter of 1985. At the end of 1 month, 31 surveys were returned (97%) and constituted the sample used in this study.

The results of the teachers' perceptions regarding the five microcomputer courses are found in Chart 2.

From the 1985 (second) survey, it was found that the teachers, who had finished 18 hours of computer courses, liked the elementary courses very much (94%). The majority felt all teachers should be enrolled in a 2-hour beginning course (84%), although the remaining 16% were less sure teachers needed a course.

Seventy-eight percent of the teachers felt they developed programs which they could use in their classrooms, while 21% did not. The general ratings of the teachers for the computer science courses were 48% positive and 52% negative.

The results provided from the survey answered the purposes of this second study, as they gave evidence of the attitudes, concerns, and opinions from these 31 students. Positive attitudes were expressed concerning the elementary

Chart 2

Teacher Perceptions Regarding Five Microcomputer Courses

Reflecting on the sequence of microcomputer courses you had at Ball State University, how would you answer the following?

1. The Elementary courses (598, 560, 590) were beneficial to my classroom teaching.

Very much				Not at all
77%	17%	3%	3%	

Comments:

2. In my opinion all teachers should enroll in the beginning course (ED EL 598).

Very much				Not at all
74%	10%	6%		10%

Comments:

3. Did you develop some programs you could use in your classroom?

Very much				Not at all
59%	19%	19%	3%	

Comments:

4. My general attitude toward the elementary education classes:

Positive				Negative
59%	35%	3%	3%	

Comments:

5. My general attitude toward the computer science classes:

Positive				Negative
25%	23%	10%	17%	25%

Comments:

Narrative:

6. What about the classes that were least appealing and rewarding to you?

   Responses: CS classes; magazine readings, graphics

7. What about the classes that were most appealing and rewarding to you?

Responses:	Practical classroom applications; time to work on own program; enthusiasm and patience of instructors, the practicum; reviewing software; developing computer curriculum; sharing ideas, readings in magazines and other literature; assignments incorporated into classroom curriculum.

8. What changes would you suggest?

Responses:	More classes from EDEL; helpful hints that make programming easier; a class on LOGO PASCAL; information through a modem; eliminate computer science courses; data base/ spread sheets, several professors for each night, more workshops with sharing ideas; more emphasis on using the computer in the problem solving process, more attention on how to use software in the classroom

---

microcomputer courses, and they had used their own original microcomputer programs in the classroom. Some concerns and opinions were directed toward the computer science courses, graphics, and readings. They listed a variety of opinions of changes or additions to the program. The teachers' perceptions and attitudes toward elementary education microcomputer courses at Ball State were positive.

Results of Both Studies

The teachers in these two studies were enrolled in a program for a graduate level teaching endorsement involving 18 quarter hours of microcomputer courses. Surveys were given at the end of the first course and then at the end of the final course. The findings of these studies showed that the perceptions of teachers and attitudes toward microcomputers and classes were positive.

Positive responses of teachers indicate that a program can be designed and implemented to help experienced teachers to accept microcomputers, do some programming and incorporate their use as another tool in learning.

The majority of students recommended at least one course for other teachers. The implication is to continue offering courses to help a larger range of teachers in the field to deal with microcomputers in education as they prepare today's children and youth for the world of tomorrow.

EVALUATION OF VARIATIONS AMONG READING IMPROVEMENT STUDENTS THROUGH THREE DECADES

MARY HARSHBARGER
Ball State University

The establishment of a college reading program implies that the institution in question has assumed some responsibility for assisting students to make an adequate adjustment to college work. While colleges like to project the image of seeking to identify a student's potential, the actual fact is that most colleges move to incorporate reading improvement classes within the curriculum because of the desirability of raising the retention rates of those admitted. The college, by accepting students whose verbal abilities are markedly below those of the traditional college student, expects that with a reading course to provide training and some time devoted to the assimilation and application of skills, many of these high-risk students will be competent to carry a regular college program.

Ball State University followed a rather typical pattern in terms of the way it handled the issue of reading improvement for its incoming students. During the early sixties, a reading efficiency lab existed whose clients were volunteers. But in the late sixties (characterized by a concern for the education of minority and "disadvantaged youths"), the University experimented with a modified freshman year which had as a major component, a reading improvement course. Since pre- and post-testing using the Diagnostic Reading Test, Survey Section: Upper Level (Committee on Diagnostic Reading Tests, Inc., N.Y. 1952) had been begun in the Efficiency Lab, that pattern was continued in the reading improvement course--and has been maintained to the present. These pre- and post-test scores for a twenty year period would seem to offer the opportunity to examine rather objectively the effect of the course.

When establishing the reading improvement course at Ball State, the research on programs at institutions which had established such courses earlier in the decade was consulted. Some of the principles which had been found to work at these were incorporated into the course at BSU. (Williams, 1969; Ratekin, 1971; Geerlofs and Kling, 1968; Marani and Maxwell, 1968).

1. Many programs for "high-risk" students encouraged first year students to have at least one meeting a week with a counselor.

2. Courses for such students usually carried credit. It was agreed that these students needed to see tangible evidence of their academic success quickly, since establishing and meeting long-term goals had not been part of their previous pattern of behavior.

3. The program for these students should be personalized, since success depends on the motivation of the individual to a great degree.

4. To be successful, such programs must have the full support of top level administrators.

5. High-risk students respond best to systematic, day after day direct teaching.

6. Colleges with programs for "high-risk" students reported that the typical student in such programs used inefficient methods to obtain ideas from print. The student was frequently a word-by-word reader who did not understand the purposes for reading assignments. Students also had difficulty in deciding what is important and what is non-essential.

Based on the reports from programs at other institutions, the reading course at BSU included the following characteristics:

1. A modified contract plan was utilized to meet the special needs of the individual student.

2. The contract plan allowed for a conference with the instructor once a week. This gave the student feed-back on his/her progress and also provided opportunities to discuss problems in other academic areas and frequently personal problems.

3. The course was a credit course. The four quarter hour course counted toward no major, but did count in terms of hours completed. The course met every day.

4. The course emphasized comprehension skills, techniques for study and textbook reading.

5. The director of the experimental program and the first two directors of the Academic Opportunity Program (AOP) viewed the reading course as vital to the success of their efforts. While monetary support was limited, other decisions demonstrated this essential support.

At the inception of the reading improvement course, a single instructor taught all sections, developed materials and did the counseling. Later, a doctoral candidate served as a second instructor. In 1971, a study undertaken by the administration concluded that the Directed Admissions Students Experimental Program (DASEP) had achieved its goal of increasing the retention rate of high-risk students and identified the reading course as a major contributor to the success of that program. It was then recommended that all "students on warning" (i.e. those students who had an SAT verbal below 360, or had graduated in the lower half of their class), take the reading class as part of a program entitled the Academic Opportunity Program (AOP).

A decision by the University Senate required all freshmen meeting the above criteria to take the reading improvement course. It became necessary to staff from 9-13 sections of the reading improvement course a quarter. The instructor's role changed to that of coordinator and consisted of working with graduate assistants, doctoral fellows and temporary faculty. While the individual styles and methods of these instructors differed greatly, overall BSU was fortunate to find caring individuals who worked hard to help the underprepared students.

Longitudinal studies of students continuing through sophomore, junior and senior years revealed that 62% of AOP students graduated in five years, 20% transferred to other institutions, 12% dropped out, and 6% experienced academic dismissal from Ball State University. These studies indicate that both during and after the one year on the program, the average scholastic performance of these high-risk students was comparable to that of the University average even when such variables as sex and class were considered. (Faculty Letter, BSU, October 16, 1978).

For the purpose of this study, it was determined rather arbitrarily to report test data from the Fall Quarter of five separate years that were assumed to be representative of the various stages of the program. In an attempt to keep the data comparable, the test scores reported are consistently from the nine and ten o'clock sections of the Fall Quarter sections of the reading improvement course.

The following table describes the sampled populations.

TABLE 1

YEARS AND DESCRIPTION OF STUDENT POPULATIONS IN THE SAMPLE FOR REPORT

Years	Description
1964	Volunteer Group (Reading Efficiency Lab)
1968	First Experimental Group (DASEP)
1974	AOP Group
1978	AOP Group
1984	First AOP Group after mandated changes: credit/no credit class; renumbered to indicate below college level work.

It would seem appropriate at this point to describe the five sampled groups more precisely. The 1964 volunteer group consisted of regularly admitted BSU students who either recognized or had an instructor who recognized that some assistance in reading would be beneficial to the specific individual. The group displayed sufficient self-discipline and motivation to continue to attend a non-credit lab for at least one academic quarter. Examination of the percentile ranks (Table 2) reveals a wide-range of abilities. Thirty percent of the group scored below the 10th percentile upon entrance, but approximately 12% entered the lab with scores above the 50th percentile.

Actually, the experimental 1968 group more closely resembled the volunteer group than it resembles the AOP groups which followed it. For this group of directed admission students (DASEP), a "self-selective" element was undoubtedly in operation. While it was expected that the experimental DASEP program would attract serious, motivated students, draft deferment may have also played a part in the motivation of some of the male students. An examination of Table 2 indicates that 36 percent of this group began the program with reading scores below the 10th percentile, and that only 4 percent of the group earned scores above the 50th percentile.

After the establishment of the AOP Program, large numbers of students came to the reading course with extremely limited reading skills. Each year a number of students were enrolled in the course whose scores did not even place them on the percentile chart for entering college freshmen. In fact, Table 2 reveals that during each of the last three sampled years, approximately 65 percent of the class entered reading below the 10th percentile.

TABLE 2

PERCENTILE SCORES OF SELECTED READING GROUPS ON PRE- AND POST-TESTS USING THE DIAGNOSTIC READING TEST

Percentile Rank	1964		1968		1974		1978		1984	
	Pre-	Post	Pre-	Post	Pre-	Post	Pre-	Post	Pre-	Post
0- 5	7	2	8	2	27	16	34	20	20	8
6-10	7	4	10	4	9	10	6	10	9	15
11-15	4	4	7	3	6	3	10	6	5	4
16-20	4	5	9	6	6	12	5	7	3	5
21-30	8	5	8	10	6	4	4	8	3	7
31-40	7	7	4	9	0	1	2	8	1	1
41-50	5	5	2	8	0	3	1	2	1	0
51-60	2	6	2	4	1	1	0	0	0	2
61-70	3	6	0	2	0	4	0	1	0	0
71-80	0	2	0	1	0	1	0	0	0	0
81-90	1	1	0	0	0	0	0	0	0	0
91-99	0	1	0	1	0	0	0	0	0	0
	48	48	50	50	55	55	62	62	42	42

Table 3 reports both pre-test and post-test scores for the sample population.

TABLE 3

TOTAL ACHIEVEMENT MEANS ON DIAGNOSTIC READING TEST READING IMPROVEMENT GROUPS BALL STATE UNIVERSITY

Dates	Numbers	Pre-Test Mean	Post-Test Mean	Change
1964	48	60.83	81.75	+20.92
1968	50	59.66	68.10	+ 8.44
1974	55	49.83	58.90	+ 9.07
1978	62	53.16	56.63	+ 3.47
1984	42	48.66	53.97	+ 5.31

An examination of these means dramatically highlights the decline in reading competencies reported for high school graduates during the decade of the seventies. With the exception of the 1978 group, the change from pre-test to post-test for each group was significant at the .001 level. Since the group sampled in 1978 appears to have made less growth than usual, the writer should state that when all students taking the course in the Fall Quarter of 1978 are considered, the change was significant at the .01 level.

In spite of the size of the classes and the severity of the reading deficiencies these students displayed, the instructors in the reading improvement course continued to individualize assignments and to have weekly conferences with each student. Student evaluations have clearly and consistently indicated that the efforts of the reading faculty to meet personal needs were recognized and for the most part appreciated.

In terms of improvement, the course has been able to produce gains in what is basically a nine-week time frame. Overall the students in the lowest percentiles make the greatest gains. Of course, one could argue that they have the greatest opportunity to grow--and possibly they may be more motivated than students who read slightly better and who assume that they will "make it" anyway. The writer can't help but feel that motivation plays a role. The 1968 group was made to feel special--as an experimental group, a group which committed themselves to an entire year of attending modified courses and sessions where they interacted with advisors and peers. Certainly the improved scores for that group suggests something special was at work.

In 1981, the AOP Program which was the academic umbrella under which the reading improvement course operated was recognized by the American Association of Schools, Colleges and Universities as a program which saw "a problem, an opportunity for better service and creatively pulled together the resources for a good local solution." (J. A. Nisbet, personal communication, May 1, 1982).

Frankly, the AOP program appears to have been a victim of its own success. Shortly after receiving the award, there was a change in directors of the program. This time, the new director was selected by the administration with no input from the faculty working within the program. Key administrators stated a fear that BSU would become known as an "AOP" university. Shortly thereafter the reading improvement course was forced to change its format. Labelled a "remedial" course, renumbered to clearly indicate that it is not a college level course, it was mandated that it become a credit/no credit course which did not count in any way except in the area of tuition costs. Using this as a rationale, advisors did not recommend it. Although the University announced that higher admission standards had gone into effect, the students who enrolled presented the same weaknesses in vocabulary development and comprehension that the staff had been dealing with for fifteen years. The scores on the 1984 pre-test certainly show that these students were not better readers than students who had immediately preceded them. Motivation of students under credit/no credit has become much more difficult. Within the last year, the AOP program has been subsumed under the umbrella of a newly established University College which has been set up to meet the "needs of underprepared students and students who have not declared a major."

Thus, the future role of this class is very uncertain. Some generalizations do seem appropriate at this point:

1. A reading improvement course can survive only if it is supported by top administrators. Philosophical differences among the faculty will undoubtedly exist--consequently, top administration must favor it, if it is to operate. The fifteen year history of the course at BSU clearly illustrates this point.
2. A reading improvement course, especially if part of a planned program,can aid in the retention of "high-risk" students.
3. College credit for the reading course is important for most "high-risk" students. Motivation of these students is easier when credit is possible.
4. Advisors must accept the value of a reading improvement course.
5. Reading competencies of entering college freshmen have shown a decline over the past decade.
6. Admission officers could help "prepare" high-risk students for special requirements which may include a reading improvement course during the freshman year.

REFERENCES

Geerlofs, M. W. and M. Kling. (1968). Current Practices in College and Adult Developmental Reading Programs. Journal of Reading, 11, 517-520, 569-575.

Marani, S. D. and Maxwell, M. J. (1968). A reading and Study Skills Program for Medical Laboratory Assistants. Journal of Reading, 11, pp. 615-620.

Ratekin, N. (1971). The Effect of Two Different Reading Programs for Culturally Disadvantaged College Freshmen." U.S. Educational Resources Information Center, Document No. ED 053 867.

Williams, R. L. (1969). Gamble on High Risk Students: How It Is Working? AAUW Journal, 62, 173-177.

A COMPARISON OF PARENT AND TEACHER PERCEPTIONS OF STUDENT BEHAVIORS

JOAN WILLIAMS
Ball State University

Introduction

Teachers use a variety of instruments and means to evaluate, assess, or appraise student achievement in school, ability levels, skill needs, interests, etc. Information obtained is helpful in planning educational programs, providing appropriate materials, and in general making the child's educational experience as beneficial as possible. This is one side of the educational "coin."

There is another side to be considered. Just as teachers need to evaluate readability of books as well as student reading ability for effective program planning, so knowledge of the affective sphere, as well as the cognitive sphere, of a student's life would hold significance.

Within this framework educators have recognized the influence on students of the important others in their lives. People holding places of such importance to students include parents and teachers. Ekwall (1983) has generalized from the body of research literature: "There is a fairly high degree of relationship between teachers' and parents' expectations and students' self-concepts." Do parent and teacher perceptions about the child agree as they look toward a common goal of seeing the student succeed as a learner?

Reports of parent feelings, concerns, and attitudes toward public education have been published in commercial magazines, the public press, and educational periodicals in recent years. Granger, (1985), an author-newspaper columist and parent, shared his concerns about educational practices and policies in a Reader's Digest article. The view he expressed relative to his son's initial experiences in school and the parents' related experiences with school personnel did not agree with the view attributed to the teachers. It is evident that teachers and parents had much different perceptions of the boy's behaviors, needs, etc. Such conflicting views surely could have had an influence on the student.

Studies supporting the importance of affective factors have been made by McCafferty (1980) and Cooper (1982). Athey (1985) presented a compilation of reading research on affective factors. Contributions by Rosenthal and Jacobson Palardy, and Purkey point out that teacher expectations affect teacher behavior toward students and students' self-perceptions and achievement.

Studies of parent perceptions by Katkovsky (1964), Brook (1975) and Wolff (1966) have indicated that parent involvement in children's school-related activities is tied to parent expectations. The completion of assigned work by the student at school and at home correlates with parent expectations and involvement.

According to research findings presented by Searls, Mead, and Ward (1985) elementary students watch television extensively on a daily basis. Fifty percent of nine-year-old children (U.S.A.) view television for three or more hours daily. While parents often permit this behavior, one can question the degree to which teachers or parents approve this practice or behavior.

Phi Delta Kappa funded a 1985 study of the attitudes of the public toward the public schools. Parents perceived a need for change in school practices by calling for more homework than teachers were assigning.

Both parent and teacher perceptions may influence factors which affect a student's learning and interest in learning. Agreement between teacher and parent perceptions of student characteristics or behaviors seems desirable. Continuity of home and school effort would seem likely to be helpful to the student. This paper describes a comparison of teacher and parent perceptions, especially regarding agreement of the two groups on items related to student behavior.

Procedure and Subjects

The primary objective of this study was to examine data provided to the Ball State University Reading Clinic by home and school to identify mutual agreement and, perhaps, support by both agencies--of the child. Such support could be an important factor in a child's ability to profit from "schooling." A comparison of teacher and parent perceptions of student behaviors was made in order to determine whether these parents and teachers agreed.

Specific types of behavior were considered. Do parent and teacher agree on such questions as the degree to which a child "blames others" for his problems, is untruthful, or is easily discouraged, etc.?

This evaluation also considered the degree of agreement between information provided by the home and the school related to television viewing time, the number of different schools attended, whether the child was living with both natural parents, and the learner's response to his reading problem (still trying, given up, etc.).

Both teacher and parents marked on survey sheets their perceptions of the behavior of the child as indicated by each of the following labels: blames others, cries, daydreams, becomes discouraged, friendly, gets in fights, happy-light hearted, has to be prodded "to get things done," is not truthful, educational interest span, listens to reason, nervous-irritable, obeys, popular with pals, is dishonest, talks back, temper tantrums, timid-shy, hurting pets, thumb sucking, strong fears. Each item was marked on a four-point scale from often (1) to never (4) with allowance being made for "don't know" marking.

Records in the Reading Center are filed alphabetically by student name. The sample consisted of records taken in order beginning with the first one in the file until one hundred-five complete data-forms for school age clients were recorded.

Cross tabulation of pairs of variables was made noting teacher and parent perceptions of the child's behavior (indicated on the forms noted above). Other information was considered, such as the number of different schools which the child had attended, the number of hours of week-day television viewing (indicated by the parent), whether the child were living with both natural parents, and the child's response to his reading problem (indicated by the teacher).

For a given behavior (as "blames others," "becomes discouraged," etc.) if either teacher or parent marked "do-not-know" for a specific student, both forms for that student for that item (behavior) were omitted from the cross tabulation. Granting that a lack of knowledge by either parent or teacher might be important to the specific child, this study examined agreement and degrees of differences in perceptions.

The records studied had been submitted for students from age seven to age sixteen. Table 1 shows the distribution of the students in the sample. The mean age was 9.39, just over nine years old. Seventy-one students (68%) were boys while thirty-four (32%) were girls. This proportion is not unexpected in such clinic programs.

TABLE 1. Distribution of Students by Age

Age	Number of Students (N = 105)
7	19
8	23
9	19
10	19
11	10
12	5
13	6
14	2
15	0
16	2

Table 2 shows the number of schools attended by the students in the sample. It is interesting to note that fifty-six of the 105 students had attended only one school.

TABLE 2. Number of Different Schools Attended by Each Student

Grade	Total Number of Students	Number of Schools Attended					
		1	2	3	4	5	6
1	3	3	-	-	-	-	-
2	30	21	8	1	-	-	-
3	23	14	4	5	-	-	-
4	21	10	4	5	1	-	1
5	10	5	4	-	1	-	-
6	6	2	1	2	-	1	-
7	8	1	3	3	1	-	-
8	2	-	1	1	-	-	-
9							
10	2	-	-	1	-	-	-

Table 3. A Distribution of Parent and Teacher Responses to Learner Behaviors Plus an Indication of Frequency of Agreement at Each Response Level

Behavior/Value	Friendly			Happy			Listens to Reason			Obeys			Popular			Must Be Prodded		
	P	T	Agreement	P	T	Agreement	P	T	Agreement	P	T	Agreement	P	T	Agreement	P	T	Agreement
1 Often	53	51	28*	17	12	5	10	19	2	6	30	-	16	17	9	33	25	11
2	48	49	23	66	62	41*	64	70	42*	90	68	60*	71	60	49*	44	24	11
3	2	3	-	15	24	6	25	10	3	8	5	1	8	19	2	25	47	15
4 Never	-	-	-	1	1	-	1	1	-	-	1	-	2	1	-	2	8	-
	Cries			Blames Others			Daydreams			Discouraged			Fights			Not Truthful		
1 Often	12	1	1	9	2	-	19	21	7	56	34	25	6	5	1	5	2	-
2	41	11	7	15	12	3	39	34	15	35	39	17	31	15	4	29	17	8
3	30	41	13	46	47	26*	25	23	7	6	23	-	56	46	31*	40	37	18
4 Never	1	31	-	7	16	3	1	6	-	1	2	-	-	27	-	8	26	3
	Educational Interest Span			Nervous and Irritable			Dishonest			Talks Back			Temper Tantrums			Timid/Shy		
1 Often	6	4	1	2	-	-	-	-	-	2	-	-	-	-	-	3	3	1
2	43	40	22*	14	10	4	-	3	-	20	3	1	4	2	1	9	18	3
3	35	36	13	62	63	45*	45	27	17	64	30	18	41	16	10	53	49	29*
4 Never	10	14	4	13	18	3	30	45	19	12	65	8	33	60	26	36	31	17

*Indicates agreement higher than 50 percent.

Results

Table 3 shows the frequency with which teachers perceived students at a specific level on each specific behavior - "T". Parent perception also is indicated - "P." For each behavior the frequency of teacher and parent agreement in perception is shown.

The reader will note that fifty-three parents classified their children as very friendly. Forty-eight parents classified their children as usually friendly. Two parents classified their children as seldom friendly. Teachers classified the students: 51 very friendly, 49 usually friendly, and 3 seldom friendly.

Further examination of Table 3 indicates that the perception of teachers and parents show a moderate degree of agreement (above 50 percent.) when identifying specific students as "friendly", "happy," "popular," "nervous and irritable," or "timid and shy." Furthermore, sixty percent of students classified as "listening to reason" by parents were classified as the same way by teachers. Parents and teachers agreed closely (67.4%) regarding the classification "rarely fights." "Good attention span" might be expected to reveal diversity of opinion between parents and teachers. But on this point agreement of approximately 50% was reached.

Tables 4 through 9 show that 78 of the 105 subjects (74.3%) were living with both natural parents. Most of the students (87 of the 105 subjects) were reading at primary level but 75 of the subjects (71.4%) were perceived by teachers as "trying hard." Parents of only 37 subjects (35.2%) indicated that a member of the child's immediate family had a similar educational problem and only 12 students had chores/jobs.

Forty-nine students watched two hours of television each weekday. Over 40 percent watched three hours or more of T.V. daily.

TABLE 4. Children Living with Both Natural Parents

	Number of Students	Percentage
With both parents	78	74.3
Not with both parents	27	25.7

Table 5. Teacher Perception of How Student Responds to His "Reading Problem"

Response	Number of Students	Percentage
1. Tries hard	75	71.4
2. Gives up	8	7.6
3. Resists help	6	5.7
4. Needs constant individual help	16	15.2

Table 6. Reading Level for Instruction as Given by Teacher

Grade	Number of Students	Percentage
Readiness	1	1.0
1	40	38.1
2	28	26.7
3	18	17.1
4	13	12.4
5	3	2.9
6	2	1.9

Table 7. Other Family Member (Father, Mother, Sibling) Also Has/Had Learning Problem

	Number of Students	Percentage
Yes	37	35.2
No	68	64.8

TABLE 8. Child Has Chores/Job (Stated by Parent)

	Number of Students	Percentage
Yes	12	11.4
No	93	88.6

TABLE 9. Weekday Hours of Television Viewing

Number Hours of TV	Number of Students	Percentage
0	1	1.0
1	12	11.4
2	49	46.7
3	28	26.7
4	13	12.4
5	1	1.0
6	1	1.0

Note: Over 40 percent watched three hours or more

Discussion

Parents and teachers had over 50 percent agreement regarding the following behaviors: very friendly, usually happy, usually listens to reason, usually obeys, usually popular, once in awhile blames others, rarely fights, good educational attention span, once in awhile nervous and irritable, once in awhile timid and shy. This means that considering the students classified at a certain level for a behavior, there was an overlap between parents and teachers of more than 50 percent. However, the actual numbers of students on which teachers and parents agreed were not high. On only "usually obeys" did the actual number of students agreed upon by parents and teachers exceed half of the total number of subjects.

Agreement of teacher and parent perceptions of student behavior seemed quite low for individuals. Different situations may well elicit or stimulate different behavioral responses. These subjects were applying for remedial help and may be atypical of the general population of school children.

However, the current trend toward closer communication and cooperation of home and school, involvement of parents in school classrooms and programs should help parents and teachers have similar perceptions related to students and to their education.

Some interesting points of information about these specific subjects include the following observations:

1. Although it might be expected that many of the children with reading difficulties would have a mother, father, brother, or sister with an educational difficulty also, the proportion of the subjects with such a reported family member was quite low (35.2%). Almost 65 percent of the subjects were reported to have no such family member.

2. Few students were reported to have chores or jobs. (Seven of the twelve who did have chores or jobs were eight or nine years old.)

3. Of the thirty subjects who were reported by teachers either to resist reading help, to give up, or to work only with individual help, only three students were reported to have chores or other work (10%). It appears reasonable to expect children who learn to follow-through with any work to completion to be more persistent in staying on task to finish reading assignments too. However, there was very nearly the same percentage (12%) of students who had chores or other work in the group (75) reported by teachers as still "trying hard" to learn to read. (There seemed little difference between proportion of students with chores and those without chores regarding whether they needed prodding to complete tasks.)

4. The 75 out of 105 students needing reading help whom teachers report as "still trying hard to learn" is a positive point.

5. Surveys have indicated considerable television viewing among students with reading limitations. Forty percent of these subjects were reported to watch television three hours or more during weekdays. Forty-nine students (46.7%) viewed television two hours daily. Wouldn't it be progress if teachers and parents could communicate well enough so that the students would have even thirty or forty-five minutes at home each day to read for fun! What wonders might be accomplished!

6. Teacher and parent did not often indicate the same perception of the specific child. Of course, people may behave differently in different contexts or settings. A child may behave and work in one fashion in the classroom and in another manner at home. One may be a more highly structured situation; the other more flexible. One setting may have more pressure or stress for the child than the other. Feelings for family members or peer group may enter into the matter.

Perhaps more emphasis on home and school visitations could be helpful for parents and teachers to learn more about how the child behaves "in the other setting." The trend of parents becoming involved in the child's classroom could be quite useful for better home-school communication. With improved communication the public image of the schools could be improved, and it would be interesting to see whether students would experience increased academic growth.

REFERENCES

Athey, I. (1985) Reading Research in the Affective Domain. In H. Singer and R. Ruddell (Ed.) Theoretical Models and Processes of Reading (3rd ed.). Newark, Del.: International Reading Association.

Brook, J.S.; Whiteman, M.; Peisach, E., and Deutsch, M. (1974). Aspiration Levels of and for Children: Age, Sex, Race, and Socioeconomic Correlates. Journal of Genetic Psychology 124, 3-16.

Cooper , H.; Findley, M.; and Good, T. (1982). Relations Between Student Achievement and Various Indexes of Teacher Expectations. Journal of Educational Psychology 74, 577-79.

Ekwall, E. E., and Shanker, J. L. (1983). Diagnosis and Remediation of the Disabled Reader (2nd ed.). Boston: Allyn and Bacon, Inc., p. 271.

Gallup, G. (1983). The 15th Annual Gallup Poll of the Public's Attitudes Toward the Public Schools. Phi Delta Kappan 65, 33-47.

Granger, B. (1985). Alec's Song. Reader's Digest. pp. 33-39.

Katkovsky, W. A.; Crandall, V. C.; and Preston, A. (1964). Parent Attitudes Toward Their Personal Achievements and Toward the Achievement Behavior of Their Children. Journal of Genetic Psychology 104, 67-82.

McCafferty, W. D. (1980). Personal-Social Influences in the Classroom. Education 100, 214-222.

Searls, D. T.; Mead, N. A.; and Ward, B. (1985). The Relationship of Students' Reading Skills to TV Watching, Leisure Time Reading, and Homework. Journal of Reading 29, 158-162.

Wolff, R. (1966). The Measurement of Environment. In A. Anastasi (Ed.), Testing Problems in Perspective. Washington, D.C.: American Council on Education.

EVALUATING PERCEPTIONS OF LEARNERS AT DIFFERENT LEVELS

DONAVON LUMPKIN
Ball State University

Perceptions of learners held by teachers may have significant impact on the direction and degree of pupil performance in a wide array of functions. Since teacher perceptions cover a broad spectrum and tend to vary with learner age and grade levels, it seems appropriate to carefully evaluate these views.

Prime areas influenced by teacher perceptions appear to be learner perception of self and others or self-concept. Evidence confirms that the concept of self held by an individual becomes a major factor in behavior, performance, and relationships. Educators generally agree that self-concept wields a dynamic influence on learner achievement in school. Research confirms this impact in all areas of the curriculum with evidence particularly strong in the area of reading.

Self-Perceptions

Self-perception may be viewed as the product of myriad experiences and interrelationships of each individual--at home, in school, and through the total environment. Most definitive evidence of influence has been found to be exerted by "significant others." In the educational setting teachers and peers rank among these "significant others."

Parallel influences on perceptions which have been subjected to less precise research could be designated as social or cultural expectations. In some neighborhoods, communities, areas, or even ethnic groups, a set of criteria for general behavior and demeanor can be clearly identified. The pattern may be understood by children and youth and absorbed into self-concept as determiners of school behavior or street conduct.

Ciani (1981) has declared that "the child comes to feel about himself as he thinks others feel about him. He identifies self in terms of the appraisal of others." Jersild (1954) notes that concept of self will be influenced by examples around one, including teachers, known public figures, and perhaps in subtle ways, by models found in reading of literature, history, or biography. Other input comes through impressions gained by way of media such as television, movies, radio programs, and the like.

Teacher Perceptions

Teacher perceptions of self and of those they teach, the children and youth in their classrooms, are undoubtedly a composite of their own prior experiences in family, school, and social situations. Additional input would be received from multiple media in today's world. Since perceptions of self and others seem to form a core from which behaviors radiate and messages are transmitted, these perceptions take on crucial importance in the teacher-learner transaction.

Both teachers and learners hold a wide array of perceptions. Teachers, in the role of "significant others," recognize that their perceptions potentially influence classroom performance in terms of typical behavior, responsiveness, achievement in general and in specific subject areas, as well as participation in the total school program. Thus, although a classroom group may be composed of a wide variety of self-perceptions, that group would be dynamically affected by the way they are viewed by their teacher.

Assessment of Perceptions

Assessment of perceptions held by teachers and implications thereof require data regarding the variety and range of views held of learners. A project to help ascertain teacher perceptions was designed to secure responses from teachers by requesting that they provide descriptors of pupils at various grade levels in school. Specifically, a sample population was asked to describe learners at three widely separated grade levels in school: Grades 1, 5, and 10. Participants, including teachers and prospective teachers, were asked to record a word or phrase which they felt described "a youngster in grade one." After a brief interval, they were asked to describe "a pupil in grade five." Finally, they were requested to write a descriptor for "a student in grade ten." In addition, teachers were asked to indicate the grade level of their current assignment, and prospective teachers recorded their preferred level for a teaching assignment.

The raw responses became the database of descriptors for learners at each grade level. Each descriptor was then subjected to judgment of whether it revealed Positive, Negative, or Neutral perception of learners at that level.

Sample responses judged positive for grade 1 included "curious," "eager," and "enthusiastic." Negative value was assigned to responses like "afraid," "anxious," and "uncoordinated." Examples of descriptors deemed Neutral were "beginning," "innocent," and "young."

For grade 5, samples of descriptors judged as Positive were "interested," "growing," and "self-confident." Those judged Negative included "apathetic," "critical," and "know-it-all." Neutral was assigned to "busy," "careful," and "changing."

At grade 10 examples of Positive descriptions were "determined," "maturing," and "thinking." Negative value was given to "bored," "insecure," and "rebel." Neutral examples were "half adult," "follows trends," and "driver's license."

Analyzing Data on Teacher Perceptions

Table 1 shows the number of descriptors that were assigned to each category from a group of 53 teachers and 47 prospective teachers. This constitutes a subgroup of 100 of a larger sample whose responses were recorded in the Spring of 1985, and the group for which statistical analysis has been completed.

The steady decrease in numbers of descriptors that were deemed positive is readily observed. This is, of course, paired with a steady increase in the negative descriptors from grade 1 to 5 to 10. Computation of chi-square reveals that differences in these responses are statistically significant beyond the .01 level of probability.

A question could be raised that using a specific grade level rather than designating primary, intermediate, or secondary student could result in teachers at that level providing a disproportionate number of positive responses. To check on this possibility those who designated grade 5 as their assignment preference were checked. (N = 14) Descriptors of grade 1 pupils were rated positive for 8 and for grade 5 by only 3; these were consistent with proportions for the total group of 100 with no evidence of partiality for their own grade level.

Impact of Perceptions on Learners

Evidence that perceptions of learners held by teachers become less positive and increase in negative valence as children progress through school is a matter for thoughtful consideration. The nature of contacts with others in their world determines to a large extent how individuals see themselves. Within the school environment, interaction with teachers would strongly influence perception of self as strong or weak, wanted or unwanted, able or inadequate, bright or dull, etc. Since learners tend to see themselves in terms of teacher expectations, the view of that teacher can be a strong factor in eliciting different behavioral characteristics and academic accomplishments.

Table 1

Number of Descriptors of Learners in Grades 1, 5, and 10 by 53 Teachers and 47 Prospective Teachers Rated as Positive, Negative, and Neutral

	Perceptions			
	Positive	Neutral	Negative	
Grade 1				
Teachers	40	5	8	
Prospective Teachers	31	8	8	
Total	71	13	16	$x^2 = 17.85$ $p<.01$
Grade 5				
Teachers	17	21	15	
Prospective Teachers	15	14	18	
Total	32	35	33	$x^2 = 12.92$ $p<.01$
Grade 10				
Teachers	8	25	20	
Prospective Teachers	4	12	31	
Total	12	37	51	$x^2 = 57.86$ $p<.01$

If teachers perceive youngsters progressing through the grades as exhibiting less interest in learning and more negative behaviors, this affects their perceptions of self and of school. Maladaptive, low-achieving learners may present aggressive, striving or withdrawn behavior patterns to maintain a concept of self consistent with negative teacher perceptions. In such a setting efforts should be made to work out ways for both teachers and learners to develop insights into the basis for their behavior.

Schools across the nation need to launch an insightful attack upon problems of increase in maladaptive behavior as children and youth progress through school exhibiting a diminishing pattern of achievement and counterproductive behavior. When pupils feel they are perceived and evaluated in an increasingly less positive manner, rapport with educational and social expectations become minimized; academic performance suffers.

Today's schools should examine ways to open up vast new domains for self-discovery and for realization of potentials. Deep and tenacious factors are rooted in the family and influences outside the school, but individuals can still change, and schools continue to be viewed as change agents.

Jersild (1954) asserts that the school has countless opportunities to help in establishment of a self-system. In school, individuals discover aptitudes and abilities, face difficulties, realize that situations exist which limit what one can and cannot do. Attitudes regarding worth are often affected since life at school is heavily invested with praise-blame, pride-shame, acceptance-rejection, success-failure!

Wlodkowski (1984) reminds educators of evidence which suggests that the way learners feel about themselves and their ability to do school work is related to what they think others expect of them. For example, students hold low academic self-concepts when they perceive parents or teachers as having low expectations of them. Thus, success of students in classes seems more closely correlated with perceptions of what others expect of them than casual observers might think.

Everything that enters into the relationship between teacher and learner appears to have some significant bearing on the learner. Teachers are not mere purveyors of knowledge, they should be openers of doors, awakeners to reality! Prescott (1957) has long urged that teachers provide experiences and environment which encourage looking ahead, believing in self, hoping, aspiring. To accomplish this with convincing style implies that the teacher perceive the learner in a positive way.

Needs Evidenced by Findings

The findings of this study call for consideration of ways to alter current teachers' perception of learners as they progress through school in a more positive direction. Educators agree that forces within the classroom, cultural pressures, and adult approval should operate to bring forth achievement to the best of an individual's ability.

A summary of concerns which cry out sharply for attention include:

1. Better awareness of the ideas which learners draw from teachers in evaluating themselves. Such understandings might open avenues for improving adaptation to the school life situation for all involved in the educational experience.

2. More adequate understanding of class groups, individual factors that influence behavior and role of teachers in enhancing learning. These insights can help to explain how teacher perceptions and actions affect group dynamics and cohesiveness.

3. Evaluating teacher perceptions and impact on learners. Analyzing views and reviewing implications can assist in modification in a positive direction.

Recommendations for Modifying Perceptions

Efforts to deal with concerns related to teacher perceptions and self-concept require early and continuing action. Care should be exercised to avoid superficial approaches or explanations of simple self-fulfilling prophecy (Rosenthal & Jacobson, 1968). The aim should be to obtain some basic and significant changes which inculcate teacher perceptions that lead to reorganization of behaviors toward more effective and productive concepts. A strong contribution to excellence in education may be achieved through strengthening of appropriate, realistic teacher perceptions.

The following recommendations are suggested for developing positive self-perceptions of learners to help strengthen positive perceptions and self-esteem of teachers.

1. Provide reading learning situations which foster experiences of success (by teachers).

2. Accept individuals and encourage uniqueness rather than uniformity (among teachers and learners).

3. Utilize bibliotherapeutic procedures and reading materials which deal with personal concerns (of teaching and learning).

4. Expand opportunities for developing and reading about particular preferences (of self and others).

5. Assist individuals in evaluating progress toward goals (and positive perceptions).

6. Strive to incorporate into the reading program components which enhance feelings of self-worth and dignity (for the entire school).

References

Ciani, A. J. (1981). Motivating reluctant readers. Newark, DE: International Reading Association.

Jersild, A. T. (1954). "Understanding others through facing ourselves." Childhood Education, 30, 411-414.

Lumpkin, D. D. (1984). Self concept and reading. In R. A. Thompson & L. Smith (Eds.), Reading research review. Afton, MN: Alpha Editions, Burgess Publishing Co.

Prescott, D. A. (1957). The child in the education process. New York: McGraw Hill Book Co.

Rosenthal, R., & Jacobson, L. (1968). Pygmalion in the classroom. New York: Holt, Rinehart and Winston, Inc.

Wlodkowski, R. J. (1984). Motivation and teaching: A Practical guide. Washington, DC: National Education Association.

RESEARCH IN TEACHING CONTEXT CLUES: AN ANALYSIS OF BASAL READING SERIES AND READING METHODOLOGY TEXTBOOKS INSTRUCTIONAL IMPLICATIONS

JOAN ELLIOTT
Indiana University of Pennsylvania

INTRODUCTION

The following paragraph represents a typical context clue review lesson that can be found in most basal readers.

> Remind pupils that they can often learn the meaning of an unfamiliar word in a sentence by looking carefully at the rest of the sentence. If there are not enough clues in that sentence, they should look at the other sentences in the paragraph, both those preceding and following the sentence with the unfamiliar word. Point out that the context sometimes gives an exact definition, but usually gives only clues that pupils must put together to find the unfamiliar word's meaning (Ginn and Company - Level 12, 1984).

In reviewing the methodology books for this study, the author retrieved a text of Paul McKee (1948) from the dusty bookcase and found that the following had been underlined as an important idea to remember:

> It is indeed unfortunate that, although teachers admonish pupils to use the context to find the meaning of a strange word or group of words, very few children receive any definite instruction in how to use the context. The result is that many of our pupils and students have little if any skill in using this important tool (p. 73).

Almost all textbooks used in reading method courses have some statement concerning the use or importance of context clues and reading. The following are representative:

> Context clues are perhaps the most important single aid to word perception (Gray, 1960, p. 25).

> It would be difficult to overestimate the value of the context in children's word perception (Cordts, 1965, p. 16).

> Context clues can help children make educated guesses about the identity of unfamiliar words (Burns and Roe, 1980, p. 117).

> Context analysis is probably the major tool for building a semantic vocabulary (May, 1982, p. 55).

> As children progress in reading, there are more and more occasions where physical context and pictures are not very useful, and they need to rely on the language of the text to help identify words they do not immediately recognize (Finn, 1985, p. 128).

The above represents some historical and current statements concerning the usefulness of context clues. These statements indicate that the use of context clues have long been considered as being very valuable in aiding a child to recognize words and as an aid in getting meaning. The instruction in using context however has been lacking in many classrooms. Teachers have only told children to try to "guess" what the unfamiliar word might be from the meaning of the sentence. As a result of this practice many children have not developed the technique of using context.

Professors of reading in teacher preparation programs (including this writer) have introduced and tested students on knowledge of specific context clues. Examination of context clues has often been based on a classification suggested by McCullough (1943, 1945). Professors of education are frequently urged to examine what they are teaching in light of what is being taught in schools. This is not designed to secure a combination of ideas from teacher educators and from practicing teachers hoping to get the best of both "worlds."

PURPOSE OF STUDY

The major purpose of the study was to determine if the various classifications and definitions of context clues suggested in reading methodology texts were in fact being presented in basal reading programs. Amoriell and Hofler (1984) attempted to analyze various commercial materials that "sensitized children to semantic cues," and found "no organized or consistent development of different types of semantic cues" (p. 383). They also found that only one type of semantic clue was adequately developed, and this clue represented the experience type semantic cue. These researchers further stated that "an analysis of 1072 semantic cues in literature appropriate for elementary/middle school levels supported the existence of seven types of semantic cues" (p. 383). The following types of semantic cues were identified:

Type	Category	Percent of Occurrence
I	Explanation/Appositive	13%
II	Experience	24%
III	Mood/Tone	9%
IV	Example	7%
V	Summary	19%
VI	Synonym	21%
VII	Comparison/	7%
	Contrast	7%

DESIGN OF THE STUDY

The first part of the study consisted of analyzing ten textbooks that are utilized for reading method courses in teacher education colleges and universities. The material related to context clues, definitions and examples could be organized into 27 categories of different context clues. The second part of the study consisted of analyzing three basal reading series to determine if the teaching activities related to context clues that were included in the teachers' manuals matched the categories found in the methodology textbooks. For example, manuals were checked to see if activities were provided specifically for the "Definition" context clue. The third part of the study sought to determine whether it was possible to identify a specific activity as a "Definition" clue if not so designated in the manual. The resulting group was developed into a list of occurrences in each category.

TEXTBOOK ANALYSIS

Ten textbooks were used for the study. These books were chosen based on the recency of copyright dates and extent of use in methods courses. The textbooks used included:

A. Teaching Them to Read, 4th edition, Dolores Durkin, Allyn and Bacon, Inc., Boston, Mass., 1983.

B. Teaching Reading to Every Child, Diane Lapp and James Flood, Macmillan Publishing Co., Inc., New York, 1978.

C. Content Area Reading, Mary M. Dupuis and Eunice N. Askow, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1982.

D. Teaching a Child to Read, Roger Farr and Nancy Roser, Harcourt, Brace, and Jovanovich, Inc., New York, 1979.

E. Elementary Reading: A Comprehensive Approach, John N. Manfieri and Lois A. Bader, McGraw-Hill Book Company, New York, 1980.

F. Teaching Reading, J. Estill Alexander et. al. Little, Brown and Co., Boston, Mass., 1979.

G. Teaching Reading in Today's Elementary Schools, Burns, Roe and Ross, Houghton-Mifflin Co., Boston, Mass., 1984.

H. Reading As Communication, Frank B. May, Charles Merrill, Columbus, Ohio, 1982.

I. The Teaching of Reading, Martha Dallman et. al. CBS College Publishing, Holt, Rinehart, and Winston, 1982.

J. Developmental Reading, K-8. Teaching from a Psycholinguistic Perspective, Daniel Hittleman, 2nd ed., Houghton-Mifflin Co., 1983.

Texts A and F suggested teaching McCullough's seven classifications of context clues. Table 1 presents the twenty-seven categories of context clues, definitions and examples found in the ten textbooks. Table 2 presents the twenty-seven categories of context clues indicating which textbook included these context clues. For example, the context clue "Definition" was suggested by all textbook authors and the context clue "Examples" was suggested by authors C, F and I.

An examination of Table 2 reveals that the authors of the ten methodology texts are not in total agreement regarding context clues to include. All the authors imply that context clues are useful but few give concrete examples for teaching and some authors don't define context clues. The first intention in this study was to compile a list of context clues according to the various authors. A review of Table 2 will indicate the context clues and the number of authors who include that category of context clues.

BASAL READER ANALYSIS

Three Basal Reading Series K-6 were used for the second part of this study. They are D. C. Heath's American Readers (1983), Ginn and Company (1984), and Scott, Foresman (1983). The objective in analyzing the reader was to determine if there was a match with the category list generated from the methodology textbooks and then

TABLE 1

CONTEXT CLUES, DEFINITIONS AND EXAMPLES

1. Definition - To introduce and define the words within the same context.
   Ex. The cornea is the transparent outer coating of the eyeball.

2. Experience - To use past experiences to complete a thought.
   Ex. Jack gave his dog a ________ to chew.

3. Comparison or Contrast - The unknown word is compared to something known.
   Ex. You do not have to run, you can ________.

4. Synonym - The use of words with similar meanings within the same text.
   Ex. I hanker to visit California. This restless longing never ends.

5. Familiar expression - The word is recognized by its use in a familiar language pattern or verbal expression.
   Ex. "See ya later, alligator!"

6. Summary - A detailed account of the happening is given. (Grouping various previous concepts.)
   Ex. Mary is a student who is interested in intellectual inquiry as well as research. She has a curious creative mind. Mary is believed by many to be inquisitive.

7. Reflection of a Mood or Situation - The tone or mood of a sentence suggests the nature of the new word.
   Ex. The old cellar was dim and cool and quiet. After being out in the sun all day, it was a relief to go down there.

8. Simile - The words as and like are used to introduce a simile.

   Ex. Her hair was as blond as harvest wheat.

9. Apposition - The meaning of a word is expressed directly after word and is set off by commas.
   Ex. The gully, a deep ditch, was full of water.

10. Antonym - a word that is opposite in meaning from another.
    Ex. Yesterday Ann was unhappy with her job, but today she is elated.

11. Groupings - An unfamiliar word is classified in context with other words that depict similar meanings.
    Ex. His tie was a combination of many colors: yellow, red, and chartreuse.

12. Picture clues - Fine line sketches are drawn to illustrate sentence or word meaning.
    (picture of truck)
    Ex. Annie drove the ________ down the highway.

13. Lexical clues - A clue that relates to the vocabulary of a language as distinguished from its grammar and construction.

    (bottle)
    Ex. _______, the rich and creamy white substance, is made of egg yolks and oil.

14. Relational clues - The words in the context relate to the missing word.
    Ex. The girl in the green bathing suit dove into the _______ with her arms stretched above her head.

15. Interpretative clues - To use the context of the sentence to interpret meaning.
    Ex. The dazzling light was like a ________ in the darkened sky.

16. Semantic clues - A package of clues provide meaning to the unfamiliar word.
    Ex. Emily Bronte grew up on the English moors. There were rolling hills covered with little shurbs and clumps of long grass. The ground was soggy and the air was damp.

17. Inferences - To piece together the meaning of a word without a literal definition.
    Ex. Peter got up early and knew it would be a glorious day. Sunlight played on the lawn while white clouds puffed by. The birds were singing like crazy.

18. Figures of Speech - To compare a familiar part of a phrase with the less familiar.
    Ex. The puppet moved like a tiny man, old and bent.

19. Idioms - words or phrases that do not mean what they say on the surface; they cannot be literally translated.
    Ex. The neighbors had fought for years over small things. One day they talked it over and decided to bury the hatchet.

20. Cause and Effect - One part (either cause or effect) is familiar to the child. If the reader has experience with one part, the other can be discovered.
    Ex. Billy was late getting home for dinner again. He knew he might be punished.

21. Syntactic clues - The way that the words being used in certain ways mean in relation to each other.
    Ex. Jenny was able to ride her tricycle much faster than her brother.

22. Example clues - Similar to a simile but more explicit.
    Ex. The letters "un" are a common prefix in our language, as in unhappy, unclean, and untrue.

23. Building Block clues - These are simply derivatives - words built by adding suffixes and other word parts.
    Ex. She unwillingly walked to school.

24. Referral clues - no definition available

25. Main idea - no definition available

26. Preposition - no definition available

27. Restatement - Sentences containing appositives, the punctuation marks of parenthesis and dashes, and the words in other words.
    Ex. The parcels all were sent down a chute - a kind of slide - into the storeroom below.

TABLE 2

CONTEXT CLUES AND METHODOLOGY TEXTBOOKS

Context Clue	Methodology Textbooks									
	A	B	C	D	E	F	G	H	I	J
DEFINITION	x	x	x	x	x	x	x	x	x	x
EXPERIENCE			x			x			x	
COMPARISON OR CONTRAST			x	x	x	x	x		x	
SYNONYM	x	x	x	x	x	x			x	
FAMILIAR EXPRESSION			x			x	x		x	
SUMMARY	x	x	x		x	x				
REFLECTION OF A MOOD OR SITUATION			x			x		x	x	
SIMILE	x	x								
APPOSITION	x	x					x		x	
ANTONYM	x	x								
GROUPINGS	x	x							x	
PICTURE CLUES		x				x	x		x	x
LEXICAL CLUES		x								
RELATIONAL CLUES		x								
INTERPRETIVE CLUES		x						x		
SEMANTIC CLUES				x						
INFERENCES				x				x		x
FIGURES OF SPEECH				x						
IDIOMS				x						
CAUSE AND EFFECT	x								x	
SYNTACTIC CLUES				x			x		x	x
EXAMPLE CLUES	x						x			x
BUILDING BLOCK CLUES								x		
REFERRAL CLUES									x	
MAIN IDEA									x	
PREPOSITION									x	
RESTATEMENT										x

to classify the various context clues included into Instruction, Review and Practice activities. Contextual activities were selected from each series and analyzed to determine the type of clue and whether the activity was instructional, review or practice. Basal series generally did not name context clue activities. Most activities were listed as, "uses context" in pictures, meaning and syntax, word referents, pronouns, appropriate wording, unfamiliar words and identifying word meaning through context clues. In categorizing the various activities with the definitions found in Table 1 many activities involved the use of several different categories of context clues. For example, a practice workbook page would have an item involving an "Inference" clue and another item where "Comparison or Contrast" clue was needed. The attempt to classify the activities into the 27 categories of specific context clues was abandoned.

Each basal series also had its own basis for what was considered instruction, practice, and review. Therefore, it decided to use the classification employed in each basal series to report the findings.

Scott, Foresman

Scott, Foresman included context clues under word identification and comprehension. Under word identification the activities were listed as: pictures, meaning and syntax. Under comprehension the activities were listed as: word referents, pronouns, adverbs, appropriate word meaning and unfamiliar words.

Table 4 presents a summary of Scott, Foresman's "uses" of context program.

TABLE 4

CONTEXTUAL ACTIVITIES: SCOTT, FORESMAN

Grades Level	Exposure	Teach	Practice	Apply	Assess	Reteach	Reinforce	** Vocabulary
K	2*							
1		22	11	5	5	6	15	
2		2	2	0	0	2	5	
3		5	30	5	4	4	6	6
4		11	25	23	10	5	8	
5		7	17	16	8	5	11	
6		6	20	23	6	8	5	
Total	2	53	105	72	33	33	50	6

* Two lessons were considered exposure prior to formal instruction.

** Six lessons were called extension of vocabulary.

Ginn and Company

Ginn and Company categorizes context clues only under vocabulary. Ginn classifies activities as I - introductory instruction; P - practice activity; R - review activity and T- skill tested. Table 5 presents a summary of Ginn and Company's context clue program.

TABLE 5

CONTEXTUAL ACTIVITIES: GINN AND COMPANY

Grade Level	Introductory Instruction	Practice	Review	Test
1	1	5	3	
2		2	2	
3	2	11	6	1
4		16	6	4
5	1	4	2	1
6	1	5	1	1
Total	5	43	20	7

D.C. Heath

D. C. Heath includes the category of context clues in the index of the manual and in their scope and sequence materials. Context clue activities were classified under vocabulary skills and in activities related to homographs, homophones and idioms. The sequence of teaching was classified as I - Introduce, R - Review, M - Maintain, and T - Test. A column titled P - Practice was included for workbook pages. Table 6 presents a summary of D. C. Heath context clues program.

TABLE 6

CONTEXTUAL ACTIVITIES: D. C. HEATH

Grade Levels	Introduce	Review	Maintain	Practice	Test
1	12	21	0	2	0
2	0	22	0	0	1
3	4	31	0	5	1
4	0	1	0	5	38
5	0	3	17	2	0
6	0	10	10	4	0
Total	16	88	27	18	40

SUMMARY AND CONCLUSIONS

The three Basal Reading Series often did not specify any particular classification of the types of context clues involved in the activities provided. D. C. Heath did teach homophones, homographs and idioms through the use of context. Many of the activities required the student to use experience to complete a thought.

A match between what methodology textbooks and basal reading series present related to context clues was not found. This study supports the findings of Amoriell and Hofler (1984) that semantic clues are prevalent in the teaching activities found in basal reading series and that the experience semantic type clue was widely developed. In quantity Scott, Foresman was far ahead with 364 contextual activities followed by D. C. Heath with 189 activities and Ginn with 75 activities. Scott, Foresman may have had more activities because their vocabulary introduction activities were categorized as vocabulary in context. Inconsistencies were evident regarding what was identified as context clue activities by the three series.

DISCUSSION

The basic concern of the study was related to classification of context clues as presented by authors of methodology texts used in teacher education and in basal reading series for children.

Teaching a classification pattern of context clues does not appear to be the answer. Requiring pre-service teachers to memorize a list of context clue categories is questionable as a way to prepare them to teach contextual analysis. Educators need to plan more creative and meaningful activities to help children to see the value of contextual analysis.

Practice could be provided in the use of syntactic or semantic clues. In the preparation of teachers, consideration could be directed to those understandings related to context clues that would be realistic and practical for application in classrooms. Authors of college methodology texts and editors of basal reader programs might work cooperatively to develop consistency in terminology and hopefully more effective strategies for understanding of context.

REFERENCES

Aaron, E.E. et al (1981). Scott, Foresman Basic Reading Program. Levels 1-11.

American Readers. D. C. Heath Company. K-6, 1983.

Amoriell, W. J. & Hofler, D. B. (1984). Vocabulary Development--Contextual and Morphemic. Reading World, 23, 383-385.

Burns, P. C. & Roe, B. (1980). Teaching Reading in Today's Elementary Schools. Boston: Houghton Mifflin Company.

Clymer, T. et al (1984). Ginn Reading Program. Levels 1-13.

Cordts, A. (1965). Phonics for the Reading Teacher. New York: Holt, Rinehart and Winston, Inc.

Finn, P. J. (1985). Helping Children Learn to Read. New York: Random House.

Gray, W. S. (1960). On Their Own in Reading. Glenview: Scott, Foresman and Company.

May, F. (1982). Reading as Communication. Columbus: Charles E. Merrill.

McCullough, C. M. (1943). Learning to Use Context Clues. Elementary English Review. 20, 140-143.

McCullough, C. M. (1945). The Recognition of Context Clues in Reading. Elementary English Review, 22, 1-5.

McKee, P. (1948). The Teaching of Reading. Cambridge: Houghton Mifflin Company.

READING INVENTORIES OF THE FUTURE

WILLIAM A. HENK
The Pennsylvania State University
at Harrisburg

Among reading educators, standardized and informal reading inventories play a critical role in diagnostic enterprises. Typically, when some confusion exists about a child's true reading ability, teachers or specialists resort to reading inventories as a more or less definitive indicator. Given the importance of these evaluative tools, it seems reasonable that every effort should be made to assure their validity and practical usefulness.

To this end, the present paper outlines recommended revisions and extensions to existing reading inventory formats in terms of (1) the nature of the materials that are included, (2) test administration procedures, and (3) the purposes to which the instruments are set. Finally, a projection is made regarding the role of a more nearly ideal use of the reading inventories of the future. While this plan is admittedly bold and ambitious, it is offered as a means of stimulating thinking which might ultimately lead to the development of a superior testing instrument.

Changes in Reading Inventory Materials

Narrative versus Expository Text. As Caldwell (1985) has suggested, distinctions need to be drawn between story discourse and exposition. Clearly these forms of text are fundamentally different in many ways and should not be mixed casually with one another. That is, it is unwise to base judgments about a child's reading ability on one type of text, without considering direct performance on the othe

Paragraph Structures. At present, little consideration is made of the various paragraph structures that occur in reading inventory passages. It may be found that a child has varying proficiencies dealing with cause-effect, comparison-contrast, simple listing, or chronological order structures. This information would be particularly useful when considered relative to specific content area texts and the proportion of various paragraph structures they contain.

Passage Length. Often an estimate of silent or oral reading ability on a reading inventory is based on a one, two, or three paragraph passage. It stands to reason that estimates of reading ability would be more reliable if the length of the passages were to be increased purposefully. Not only would increased passage length allow for expanded miscue appraisal in oral reading, but also the longer passages might more nearly reflect the macrostructures present in narrative text (e.g. setting, starter events, conflicts, rising action, climax, and resolution) as well as be more in line with the length of reading assignments given to students in the upper elementary, middle school, and junior high grades.

Survival Discourse. Current reading inventory materials tend to neglect many types of prose associated with daily living. For instance, few, if any, contain samples from newspapers, menus, contracts, information forms, computer documentation, or other procedural texts like written directions for recipes or installing household gadgets. As our society becomes increasingly technological in nature and our literacy demands change, discourse other than simple narration or exposition should be represented in the inventories.

Determinants of Prior Knowledge. Most reading inventories do not possess provisions for determining the reader's prior knowledge of passage content. Such an oversight precludes an assessment of exactly how much information a reader gains from reading. As Tuinman (1974) has noted, some children can answer a substantial number of questions on reading inventories without ever having read the passages at all. Perhaps as Johnston (1983) suggests, content specific vocabulary determinants and pre-reading/passage independent questions should be included to assess readers' prior knowledge. Such items would allow test administrators to differentiate qualitatively, for instance, between two children achieving the same instructional level when one obtains the level through information gain and the other possesses most or all of this amount of information before even reading the passage.

Variant Print Mediums. The advent of the microcomputer has brought a veritable revolution in the print mediums with which readers will have to contend. Much of the reading of the future will be made from computer terminal screens of varying resolutions, colors, and "print contrasts" (Mason, Blanchard, & Daniels, 1983). Similarly, the sizes, fonts, and other attributes of printer-generated text such as right-margin justification may require differential diagnosis as reading may be less generic, and therefore, more print-dependent than we now know or can anticipate. Reading inventories of the future will simply be more useful when performance under variant print medium conditions can be measured.

Metalinguistic Knowledge. An area virtually untapped by prevailing reading inventories is the assessment of metalinguistic knowledge (Durkin, 1978; Gillet & Temple, 1982). It would be useful to know how aware children are of graphic signals (such as headings, punctuation, italics, paragraph indentations, and formatting), as well as signal words (like first, then, therefore, for instance, in contrast, and similarly). A simple interview with children when they are confronted by unfamiliar narrative or expository text can go a long way in informally indicating how much metalinguistic knowledge a child possesses about the way text, especially exposition, is structured. The inclusion of text in which all the word identities, except signal words, are graphically obscured (by converting all other letters to x's or passing a magic marker over words) can prove to be quite revealing. This knowledge of children's linguistic conception of a word, a phrase, a syllable, a margin, and a paragraph might help the test administrator identify problems more basic than originally assumed and help explain why breakdowns in instructional interchanges between the teacher and the student may occur. It is possible that the inability of some children to understand the vocabulary used by the teacher during instruction serves as a major deterrent to learning from text.

Ideal versus Flawed Passages. Much research exists to suggest that children are expected to read an alarming amount of inconsiderate text (Armbruster & Anderson, 1981). Inconsiderate texts are often poorly organized, lacking in cohesion and transition, and are stylistically, developmentally, and conceptually inappropriate for the intended readership. Therefore, it would seem that future reading inventories should possess the capability to examine performance not only on well-structured passages, but also on less considerate ones. In fact, careful monitoring of a child's performance on inconsiderate texts might reveal otherwise elusive metacognitive strategies and coping mechanisms used by the student when faced with uncertainty.

Expanded Passage Pools. The discussion thus far points to the idea that reading inventories of the future will need to contain far more passages than are currently available in any of the existing inventories, commercial or otherwise. The passage pools should bridge diverse content areas and styles, varying readability levels, paragraph types, and discourse categories, disparate prior knowledge and interest characteristics, and progressive levels of considerate text. Ideally, all of the passages could be classified and cross-referenced in a manner similar to the Degrees of Reading Power tests (CEEB, 1982). In this way, diagnosticians will be better able to make judgments about why children perform as they do in particular basal and content area texts that they use in school as well as determine alternative materials which might prove to be instructionally more beneficial.

Changes in Procedures and Policies

Varying Reader's Purpose. Ideally, children's reading performance should be monitored under a number of possible reading purposes, and materials to allow these judgments ought to be included. Among the most important of these purposes would be to read for gist, to locate specific details, to be prepared to apply the knowledge (as in following a set of directions), to analyze critically, to respond affectively, to generalize, to retell in one's own words, and to respond to probes following reading.

Passage Availability . Since, in adult reading, it is rare that materials are ever "taken away" from a reader, diagnosticians should also be encouraged to check comprehension with the passages still available to the testee during retelling and/or questioning. The student's ability to locate specific information as well as to exercise more global searching strategies might then be discerned.

Determining Reading Potential. Presently, examiners estimate how well students ought to be able to read by reading passages to them and appraising their comprehension of the material. An interesting alternative, and one that may warrant additional research, is to evaluate how well students comprehend materials while listening and following along in a corresponding text. This practice might eliminate the decoding burden on the reader while providing a permanence of stimuli that simple listening without text availability does not permit.

Scrolling Facility. Unlike books, computer text displays "mask" available information since the screen may only show a limited portion of the total stored text. If, as Mason, et. al (1983) predict, much reading in the future will be made via CRT presentation, then one ability which should be determined is the reader's skill at locating specific text segments and important information within the electronic file, but outside the immediate screen display. And depending on the storage capacity of the disks, it may be necessary to determine how well a child can move between files on a disk or even between disks to locate information.

Process Measurements. Present reading inventories measure the products of reading rather than the process itself. Retellings and answering questions represent "after the fact" rather than dynamic, in-process assessments. In fact, the closest thing to an in-process measurement seems to be noting a child's oral reading miscues. Advances in technology may permit greater in-process measurements than were ever before possible. Perhaps the reading inventories of the future might contain instrumentation to monitor eye movement behaviors during an error detection paradigm, measure brain area activation through blood flow, analyze and classify oral reading using sophisticated voice decoders, indicate specific processing difficulties of neurological origin using the BEAM (brain electrical activity -mapping) computer (McKean, 1981), or measure immediate and long-term interest during reading through other physiological means. In a related vein, having children "think aloud" as they read might offer fascinating diagnostic information. Verbal self-report data is gaining increasing acceptance as being able to "mirror the processes of the mind" (Ericcson & Simon, 1980).

Identifying Interest. An important distinction which future reading inventories should address is between how readers fare with interesting versus uninteresting materials. This information, when coupled with prior knowledge estimates, might help improve predictions of readers' performance. Baldwin, Peleg-Bruckner, and McClintock (1985) have shown that interest and prior knowledge operate somewhat autonomously. In other words, high interest is no more a guarantee of high prior knowledge than low interest is of limited prior knowledge. Readers may be interested in a topic but have little knowledge of it; and conversely, they may know much about topics that are largely uninteresting to them. Since both prior knowledge and interest exist along disparate continuums, future assessments must consider both simultaneously and how they might interact with one another for individual readers.

A Projection for the Future

Lastly, a diagnostic setting is envisioned whereby the examiner possesses an enormous capability to retrieve and utilize text for specific reading measurement purposes. Such flexibility is within the grasp of reading educators because the majority of the technology already exists. All mainframe, most midframe, and even some hard disk-equipped microcomputers possess a storage capacity large enough to accommodate thousands of passages which could be retrieved upon command. If properly coded, passages would be identifiable to the test administrator by noting the list of accompanying descriptors. The coding, which could be similar to that used in library literature searches, would then allow for an unheard of assessment specificity.

Imagine a scene in which the examiner wants to know how a child would perform on an inconsiderate expository social studies passage of moderate interest and limited familiarity, which uses a chronological paragraph structure, when the child's eventual goal is to sequence historical events. Following the prototype outlined here, such an assessment would be more nearly possible. Most importantly, heretofore unobtainable diagnostic information would be made available to the benefit of readers, teachers, administrators and parents alike.

While the speculative scene outlined here is surely desirable, obviously much would need to be done to make such an instrument a reality. Surely, its development would depend on the collaborative efforts of numerous reading educators, computer specialists, measurement experts, and practicing school personnel. But considering the potential value of this type of instrument, any time investment made in realizing its existence would surely be worthwhile.

REFERENCES

Armbruster, B. B., & Anderson, T. H. (1981). *Content area textbooks*. (Reading Education Rep. No. 23), Champaign, Ill.: Center for the Study of Reading.

Baldwin, R. S., Peleg-Bruckner, Z., & McClintock, A. H. (1985). Effects of topic interest and prior knowledge on reading comprehension. *Reading Research Quarterly*, *20*(4), 497-504.

Caldwell, J. C. (1985). A new look at the old informal reading inventory. *Reading Teacher*, *39*(2), 168-173.

College Entrance Examination Board. (1983). Understanding the DRP: Charts and tables. New York.

Durkin, D. (1978). *Teaching them to read*. Boston: Allyn & Bacon.

Ericcson, K. A., & Simon, H. J. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215-251.

Gillet, J. W., & Temple, C. *Understanding reading problems*: *Assessment and instruction*. Boston: Little, Brown.

Johnston, P. H. (1983). *Reading comprehension assessment: A Cognitive basis*. Newark, Del.: International Reading Association.

Lassen, N. A., Ingvar, D. H., & Skinhoj, E. (1978). Brain function and blood flow. *Scientific American*, *239*(4), 62-71.

Mason, G. E., Blanchard, J. S., & Daniel, D. B. (1983). *Computer applications in reading* (2nd Ed.). Newark, Del.: International Reading Association.

McKean, K. (1981). Beaming new light on the brain. *Discover*, December.

Tuinman, J. J. (1974). Determining the passage-dependency of comprehension questions in 5 major tests. *Reading Research Quarterly*, *9*, 207-223.

REACTION PAPER

Commentary on READING INVENTORIES OF THE FUTURE by William A. Henk

Reactor: LOUISE M. TOMLINSON
University of Georgia

William A. Henk (1985) has presented very interesting and promising perspectives on the future of informal reading inventories. He addresses the current status of inventories in terms of the frequency and nature of their use, their impact on instruction, their limitations, and their liabilities. He makes some very insightful suggestions for changes in IRI materials, procedures, and purposes. He forsees a more interactive role for the diagnostician both with the computer and the individual student and, in spite of some of the imposing conditions that technogical intervention will create for society-at-large, he remains optimistic about his predictions. Henk's major emphasis is on creating a more diversified diagnostic resource in the informal inventory and amplifying its efficiency through computer-mediation.

The author's suggestions for diversifying informal inventories in terms of including more content area and survival related texts are well taken. This direction would address more than one deficiency in the current state of the art of inventories. Not only do a majority of inventories contain comprehension passages taken exclusively from basal reader-type texts, but, out of eleven informal inventories reviewed by Jongsma & Jongsma (1981), only two had a range of difficulty from pre-primer or primer to grade 12. (The more frequent cut-off point was grade 8 or 9.) Thus, not only are higher level IRI selections in scarcity, but, for the grade levels at which content area and survival related skills are most critical, appropriate inventories are practically non-existent. Henk suggests that selections of greater length be included in inventories "to increase macrostructures in narrative text and reflect longer assignments usually given to older students" -- again, a valuable suggestion since the Jongsma's review indicates that length of passages in inventories ranged from 24 to 361 words. Contrary to popular opinion, Ekwall (1983) and Valmont (1972), Henk also suggests that passage-independent questions be included in comprehension assessment to tap the reader's prior knowledge. The addition of content area and daily reading-survival discourse, as well as an increase in passage length and the inclusion of passage independent-questions, would do much to provide a more comprehensive measure of the knowledge and skills which standardized tests either do not measure or tend to overestimate.

Henk's conceptualization of the IRI as a computer-mediated instrument allows for CRT screen displays and computer printouts to be used to compare to regular paper texts for transfer effects. Reinking and Schreiner (1985) provide a relevant perspective of this notion: "As an alternative medium for presenting text the computer encompasses a unique set of technological attributes... [which may] influence cognitive processing during reading" (537). They suggest that the technological attributes of the computer can facilitate a more interactive flow of information between the reader and the text. From a study comparing groups of subjects having no option to manipulate text on paper or computer display to those with a variety of options for manipulating computerized text (requesting information to address comprehension difficulty), Reinking and Schreiner concluded that: "the computer might provide unique opportunities to manage a reader's interactions with text during independent reading" (549). Computer-mediated instructional texts which evolve from such research should, no doubt, be matched by comparable IRI formats.

Another aspect of Henk's conceptualization of technological advance in the structure of inventories is computer monitoring of oral and silent reading processes. In addition to the possibilities for charting eye movement, EEGs, blood flow, and other involuntary body functions, he also refers to the development of voice decoders. This facility would be invaluable for assessing oral reading in terms of rhythm, intonation, rate, and tone to determine independent, instructional, and frustration levels of students.

This paper presents some very exciting ideas for the diversification and computerization of informal inventories. Technology will, no doubt, aid in making inventories considerably more diversified in scope and more effective in their application. Further research on both the state of the art of inventories and the effects of computer-mediated texts would serve to provide additional direction in the development of informal inventories of the future.

REFERENCES

Ekwall, E.E. (1983). Using informal reading inventories, the cloze procedure, and the analysis of oral reading errors. In E.E. Ekwall and J.L. Shanker (Eds.), Diagnosis and Remediation of the Disabled Reader. Newton, Massachusetts: Allyn and Bacon, 366-410.

Jongsma, K.S., and Jongsma, E.A. (1980). Test Review: Commercial informal reading inventories. Reading Teacher, 34, 697-705.

Reinking, D., and Schreiner, R. (1985). The effects of computer-mediated texts on measures of reading comprehension and reading behavior. Reading Research Quarterly, 20, 536-552.

Valmont, W.J. (1972). Creating questions for informal reading inventories. Reading Teacher, 25, 509-512.

THE CLOZE READING INVENTORY AS A QUALITATIVE AND QUANTITATIVE MEASURE OF READING ABILITY

ROGER J. DE SANTI
University of New Orleans

RENEE M. CASBERGUE
Tulane University

VICKI G. SULLIVAN
Orleans Parish Public Schools

While the cloze procedure has been in existence since 1897 (Ebbinghaus), the evaluative applications of the Cloze procedure have been essentially limited to only the past two decades. The vast majority of these evaluative applications have utilized the traditional format of scoring responses to deletions as either correct or incorrect based on whether or not the response was exactly the same as the word deleted. Notwithstanding a few notable exceptions (Bormuth, 1965; Hafner, 1964; Henk, 1981; Lopardo & Bortnick, 1977; Vaughan & Meredith, 1978), such quantitative approaches have given scant attention to substantial amounts of qualitative information which can be gained from analyzing the responses which are not exactly the same as the word deleted.

Implementation of the cloze based reading inventory described herein can yield both quantitative and qualitative information by considering the semantic and syntactic acceptability of responses. Information is presented on passage construction and selection, administration, scoring, interpretation, and the validity and reliability results of having employed this technique with 864 individuals, grades 3 through 12 in 19 classrooms from two school systems.

Instrumentation

For the purposes of this cloze based reading inventory, passage construction employed the traditional format. Both the first and last sentences of the passage were left intact, and beginning somewhere within the first five words of the second sentence a word was deleted. Every subsequent fifth word was deleted and replaced with a blank of standard length until fifty deletions were replaced with blanks. In order to use a passage in its entirety, the passages selected for conversion to the cloze format were approximately 300 words in length.

Passage selection criteria

The readability level of each of the 36 passages employed in this study was initially established using an extension of the Harris-Jacobson (1980) readability formula. Each passage was subsequently checked using the Fry (1977) readability formula. Both techniques use vocabulary difficulty and sentence length as estimates of readability, although the Fry uses the average number of syllables per word and the Harris-Jacobson uses the number of words not on a given list as measures of vocabulary difficulty. Only passages which were identified by both formulas as being at the appropriate grade levels were employed. Passages, were not rewritten to satisfy readability requirements (Bruce, 1981). Passages were also compared with regard to their syntactic complexity by counting the number of sentences, t-units, and the average number of words and embeddings per t-unit (DeStefano, 1978; Hunt, 1965). Additionally, passage selection criteria provided for the exclusion of any passage which contained conversation because of its tendency to affect readability (Helm, 1973). The final set of passage selection criteria included rating the interest and appeal of each passage. Three professional reading educators used an interest checklist by grade level derived from the literature (Ashley, 1970; King, 1967; Meisel & Glass, 1970; Novell, 1958; Shepherd, 1973; Stevens, 1981). The interest and appeal of each passage was then rated by the subjects included in the field testing of the inventory. The demographics of this group are described below. A Likert type scale of one, very interesting, through three, not interesting was used. The mean grade level ratings ranged from 1.26 to 1.98 indicating that no passage was rated as being "not interesting."

Administration

The administration of the inventory included instructions, practice with a sample passage, repeating the instructions, and completing the inventory passage. The instructions to the reader included (a) read the whole story before trying to fill in the blanks, (b) write one word in each blank, (c) try to write a word in each blank, guess if you do not know, you may skip hard blanks and come back to them when you have finished the rest, (d) spelling mistakes will not be marked wrong, do the best you can, (e) please write neatly, (f) the name of this story is _______, (g) you will have enough time to finish, (h) do you have any questions, and (i) please begin.

The readers were then given an opportunity to practice the cloze procedure using a sample cloze passage which contained ten deletions rather than the lengthier fifty deletion format. The primary purpose of the practice session was to overcome the effects of unfamiliarity with the cloze procedure and eliminate interactive results stemming from test-wiseness. When the sample passage was concluded, discussion of how to determine whether or not a word is sensible for a particular deletion included the use of content and context cues, punctuation, and syntactic features such as number and voice agreement. Following the practice and discussion session, the instructions were repeated and the first inventory passage was administered. The two subsequent passages were administered in a counterbalanced format without practice passages. The amount of time required to administer each inventory passage with groups of total class size was 30 to 45 minutes.

Scoring

Having administered the inventory passage, the reader's responses to the deletions were recorded, scored, and interpreted using the cloze passage coding form illustrated in Figure 1. There are nine columns across the coding form. The first three columns provide routine information. The deletion number (DN) column identifies one line for each deleted word. The word deleted column lists each word which has been removed from the passage. The reader's choice column is where the word provided by the reader for a given blank is recorded. The remaining six columns are the coding categories which constitute the core of the scoring procedure. For each reader's choice only one of the first five coding categories is checked and the last category may or may not be checked. As a result, there will always be at least one check but never more than two checks for a given reader's choice. The blank coding category is checked if the reader supplied no word for the blank. The exact category is checked if the word chosen by the reader is the same as the word deleted regardless of spelling. The consistent and coextensive (C&C) category is checked if the reader's choice is not the same as the word deleted but is consistent with the author's message. The semantic (SEM) category credits the reader not for comprehending the author's message but for producing a statement of logical meaning. If, at this point, no category has been checked then the none category is checked. The nonstandard grammatical structure (NGS) category, the sixth and last category, is not related to the previous categories which emphasized comprehension and logical language. This category deals with the success with which the reader coped with the grammatical structures of the passage and is checked if the reader's choice is grammatically unacceptable.

As examples of implementing the scoring procedures, Figure 2 illustrates the scoring of ten different readers' responses to the same blank. After conversion to the cloze format, a sentence within a passage might read, "The __________ ate the lettuce." Reader one provided no choice for the blank, so the appropriate coding was blank. Both the second and third readers are coded as exact even though the third reader misspelled the word. The fourth reader provided a choice which is either a misspelling of the word deleted or an entirely different word. Selection of the appropriate coding category

FIGURE 1

CLOZE PASSAGE CODING FORM

Reader's Name ______________ Passage __________ Date ________

			CODING CATEGORIES					
DN	WORD DELETED	READER'S CHOICE	BLANK	EXACT	C&C	SEM	NONE	NGS
1								
2								
3								
4								
5								
--								
--								
--								
46								
47								
48								
49								
50								

TOTAL NUMBER ____________________________________

TOTAL PERCENT ____________________________________

TRADITIONAL COMPREHENSION ________

TOTAL COMPREHENSION ________

LOGICAL LANGUAGE USAGE ________

STRUCTURE OF LANGUAGE (100% - TP NGS) ________

should be based on knowledge of the reader and which of the two words was most likely the reader's intention. Readers five, six, and seven each made choices which maintained the author's message although using a word different from the one deleted. The coding is the consistent and coextensive category. The eighth reader's choice is coded semantic because, when considered outside the context of the author's intentions, the response does contribute to a logical statement. Reader nine provided a choice which neither contributed to understanding the author's message nor toward an independent logical statement and was therefore coded as none. The tenth reader's choice was coded as both none and nonstandard grammatical structure because the choice did not contribute to understanding the author's message, did not contribute toward a logical statement, and was not grammatically acceptable.

FIGURE 2

SAMPLE CODINGS

The ______________ ate the lettuce.

			CODING CATEGORIES					
READER	WORD DELETED	READER'S CHOICE	BLANK	EXACT	C&C	SEM	NONE	NGS
1	rabbit	______	X					
2	rabbit	rabbit		X				
3	rabbit	rabit		X				
4	rabbit	rabbi		?		?		
5	rabbit	bunny			X			
6	rabbit	hare			X			
7	rabbit	coney			X			
8	rabbit	horse				X		
9	rabbit	brick					X	
10	rabbit	run					X	X

Interpretation

The results of the completed scoring are used as the basis of the interpretation of reader performance. The number of checks in each coding category column is totaled and entered in the total number space at the bottom of each column. The total percent for each column is derived by multiplying the column's total number by two percent (50 deletions x 2% = 100%). The four interpretation values are then based on the total percent for each coding category. The traditional comprehension value equals the total percent of the exact coding category and is interpreted using the scale established by previous research (Bell, 1977; Bormuth, 1968). A score of 58% or more indicates the independent reading level, 44% through 56% the instructional level, and 42% or less the frustration level. The total comprehension value is derived by adding together the total percent for the exact and the consistent and coestensive meaning coding categories. This value goes beyond the previous value in attempting to identify the level of the reader's interaction with and understanding of the author's message. By combining these two, a truer picture of the reader's comprehension may be obtained. This value is interpreted using the scale established through discriminant function analysis. A score of 76% or more indicates the independent reading level, 64% through 74% the instructional level, and 62% or less the frustration level. The third intrepretative value is the logical language usage value which considers the reader's ability to generate logical language structures. This value is obtained by adding together the total percent for the exact, consistent and coextensive meaning, and the semantic coding categories. This combination of scores does not provide a measure of comprehension, which is the perspective of the two prior interpretative values, but rather a qualitative assessment of the reader's ability to make sense within the limits of our language. The fourth and final interpretative value, the structure of language value, is calculated by subtracting from 100% the total percent for the nonstandard grammatical structure coding category. The resulting value is indicative of the reader's ability to deal with syntax. This value is most useful if, within the coding category, patterns are identified. For some readers, it may be profitable to distinguish between the responses which appear to be dialectal in their origin and those which are not. Similarly, it may be beneficial to differentiate responses by readers for whom English is a second language.

Method

Subjects

The inventory selections were field tested 3,456 times with a total of 864 subjects from 19 classrooms distributed across third through twelfth grade levels in an urban and a suburban school system. The demographic characteristics of the subjects were: 434 females, 430 males, 314 blacks, 517 whites, and 33 other. The number of subjects at each grade ranged from 65 to 102.

In addition to the demographic information, two sets of data were collected for each subject. The first set included the performance on the Comprehensive Test of Basic Skills (CTBS) for the vocabulary, comprehension, and total reading sections. The CTBS was regularly administered by both school systems. The second set of data included each subject's performance on the three cloze passages for each grade level. These data were used in calculating the validity and reliability estimates of the inventory.

Data Analysis

The inventory's concurrent and predictive validity were determined by calculating Pearson product-moment correlations between the inventory's four interpretive values (traditional comprehension, total comprehension, logical language usage, structure of language) and the three CTBS scores (vocabulary, comprehension, total reading) for each of the three passages. 180 correlations (4 values x 3 scores x 3 passages x 5 grades = 180) were analyzed for concurrent validity. Because the twelfth grade (65 subjects) graduated from high school between the first and second administrations of the CTBS, this grade level was not included in the predictive validity analysis. As a result, 144 correlations (4 values x 3 scores x 3 passages x 4 grades = 144) were analyzed to estimate predictive validity. When combined across both validity analyses, the 324 correlations were distributed as 252 (77.8%) significant at the .001 level, 47 (14.5%) at the .01 level, 15 (4.6%) at the .05 level, and 10 (3%) not significant. 314 of the 324 correlations (97%) were significant at at least the .05 level.

The equivalence of forms within grade levels was established through the calculation of Omega Squared values. Of the 80 values calculated with the data from the 408 subjects in the suburban portion of the sample, 72 were below .10 and a total of 78 values were below.20. Values of .20 are considered to be quite low and values of .10 or less can be considered to be of no practical significance (Cooney & Zinkgraff, 1979; Halderson & Glasnapp, 1972). The alternate forms reliability coefficient for grades 3 through 12 was found to be .61.

Single rater reliabilities were calculated to determine if a variety of teachers could consistently score student performance on a variety of passages. From the urban sample, 10 students were randomly selected from each of the five grade levels 3, 5, 7, 9, and 11. The 150 passages of these 50 students were randomly assigned to and scored by a group of seven raters using the scoring directions. Each of the seven raters had a minimum of five years teaching experience at the elementary or secondary level and was a student in a graduate level reading specialist program. When averaged across passages and grade levels, the mean reliabilities for single-raters for the interpretive values (traditional comprehension, total comprehension, logical language usage, structure of language) were .95, .81, .86 and .58.

Inter-rater reliabilities were calculated to determine if different raters could consistently rate student performance using subjects, passages, and raters described under single rater reliability. Three raters scored each passage for a total of 450 passage ratings. Averaged across grade levels and forms, the mean reliabilities for the interpretive values were .97, .93, .96, and .75.

Results

The results of the concurrent and predictive validity studies indicate that the inventory, when compared with a standardized test such as the CTBS, is a valid measure of reading ability. The equivalence of forms analysis indicated that there is little or no practical significance in differences in student performance on the passages within each grade level, and they may be used interchangeably. As a result of the alternate forms reliability analysis, users of the inventory can be fairly certain that the use of alternate passages within a grade level will yield similar estimates of a student's reading proficiency. Following the results of the reliability analyses, the instructions for the structure of language value were refined and sample codings were developed to increase reliability (De Santi & Sullivan, 1984 & 1985). The results of the analyses and the revisions indicate that the inventory is consistent in its estimate of a reader's performance.

Discussion

By measuring reading performance through the use of cloze passages, the inventory is compatible with the perspective that reading is an interactive process between author and reader, is based in silent reading which is the predominant form of reading, and being based in connected discourse and requiring responses within the text while the text is being read, asseses reading ability while the reader is reading. As there is no information provided beyond the cloze passage itself, in the form of questions and answers, no distinctive answer may be selected and passage dependency is enhanced (Hanna & Oaster, 1978-1979; Pyrczak, 1975, 1976; Pyrczak & Axelrod, 1976; Tuinman, 1973-1974). As the cloze passages may be group or individually administered without affecting procedures or standardization, the saving in time expended can be considerable. Additional efficiency is derived from an administration period of 30 to 45 minutes and, with a minimal amount of practice, scoring in 8 to 12 minutes per reader.

By augmenting the tradition of considering only the exact response with considerations of consistency, semantic acceptability, and syntactic acceptability, the cloze procedure can be extended to applications of measuring reading achievement, determining independent, instructional, and frustration reading levels, and diagnosis of an individual's reading strengths and weaknesses.

REFERENCES

Ashley, L. F. (1970). Children's reading interests and individualized reading. Elementary English, 47, 1088-1096.

Bell, D. (1977). The emperor's new cloze. Journal of Reading, 20, 601-604.

Bormuth, J. R. (1965). Validities of grammatical and semantic classifications of cloze test scores in J. A. Figurel (Ed.), Reading and inquiry. Proceedings of the International Reading Association, 10.

Bormuth, J. R. (1968). Cloze test readability criterion reference scores. Journal of Educational Measurement, 5, 189-196.

Bruce, B. C. (1981). Plans and social actions. In R. J. Spiro, B. C. Bruce & W. F. Brewer (Eds). Theoretical issues in reading comprehension. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Cooney, J. & Zinkgraff, S. (1979). Look beyond your F-test: Estimating statistical association in ANOVA models. Paper presented at the Southwest Educational Research Association.

De Santi, R. J. & Sullivan, V. G. (1984). Inter-rater reliability of the cloze reading inventory as a qualitative measure of reading comprehension. Reading Psychology: An International Quarterly, 5, 203-208.

De Santi, R. J. & Sullivan, V. G. (1985). Reliability of single-rater judgements of semantic and syntactic classification of cloze test responses. Journal of Research and Development in Education, 18, 47-53.

DeStefano, J. S. (1978). Language, the learner and the school. New York: John Wiley and Sons.

Ebbinghaus, H. (1897). Ueber eine neue methode zur Prufung geistiger Fahigkeiten und ihre Anwendung bei Schulkindern. Zeitschrift fur Psychologie und Physiologie der Sinnesorgane, 13, 401-457.

Fry, E. (1977). Fry's readability graph: Clarifications, validity, and extension to level 17. Journal of Reading, 21, 242-252.

Hafner, L. E. (1964). Relationships of various measures to the cloze, In E. L. Thurston & L. E. Hafner (Eds.), Thirteenth Yearbook of the National Reading Conference.

Halderson, I. & Glasnapp, D. (1972). Generalized rules for calculating the magnitude of an effect in factorial and repeated measures ANOVA designs. American Educational Research Journal, 9, 301-310.

Hanna, G. S. & Oaster, T. R. (1978-1979). Toward a unified theory of content dependence. Reading Research Quarterly, 14, 226-243.

Harris, A. J. & Jacobson, M. D. (1980). The Harris-Jacobson readability formula. In A. J. Harris and E. R. Sipay, How to increase reading ability. (7th ed). New York: Longman.

Helm, E. B. (1973). Uses of the cloze procedure in elementary schools. In P. L. Nacke (Ed). Programs and practices for college reading (Vol. 2), 22nd Yearbook of the National Reading Conference. Boone, North Carolina.

Henk, W. A. (1981). Effects of modified deletion strategies and scoring procedures on cloze test performance. Journal of Reading Behavior, 13, 347-357.

Hunt, K. (1965). Grammatical structures written at three grade levels. Champaign, Illinois: National Council of Teachers of English.

King, E. M. (1967). Critical appraisal of research on children's reading interests, preferences and habits. Canadian Education and Research Digest, 7, 312-326.

Lopardo, G. S. & Bortnick, R. M. (1977). A silent reading inventory based on the cloze procedure. Illinois School Journal, 57, 19-23.

Meisel, S. & Glass, G. G. (1970). Voluntary reading interests and the interest content of basal readers. The Reading Teacher, 23, 655-659.

Novell, G. W. (1958). What boys and girls like to read. Morristown, New Jersey: Silver Burdett Company.

Pyrczak, F. (1975-1976). A response note on measures of the passage dependence of reading comprehension test items. Reading Research Quarterly, 11, 112-117.

Pyrczak, F. (1976). Context-independence of items designed to measure the ability to derive the meanings of words from their context. Educational and Psychological Measurement, 36, 919-924.

Pyrczak, F. & Axelrod, J. (1976). Determining the passage-dependence of reading comprehension exercises: A call for replication. Journal of Reading, 19, 279-283.

Shepherd, D. L. (1973). Comprehensive high school reading methods. Columbus, Ohio: Charles E. Merrill.

Stevens, K. C. (1981). Reading interests among fifth and sixth graders. The Reading Teacher, 21, 147-151.

Tuinman, J. J. (1973-1974). Determining the passage dependency of comprehension questions in five major tests. Reading Research Quarterly, 9, 206-223.

Vaughan, J. L. & Meredith, K. E. (1978). Reliability of the cloze procedure as assessments of various language elements. In P. D. Pearson & J. Hansen (Eds), Disciplined inquiry in process and practice. 27th Yearbook of the National Reading Conference.

REACTION PAPER

Commentary on THE CLOZE READING INVENTORY AS A QUALITATIVE
AND QUANTITATIVE MEASURE OF READING ABILITY
by Roger DeSanti, Renee Casberque and
Vicki Sullivan

Reactor: CAROLE F. STICE
Tennessee State University

This paper presented a description of the development and validation of a group cloze based inventory along with guidelines for a miscue - like qualitative analysis of readers' substitutions/insertions.

The test appears to have been well constructed and subjected to adequate readability, reliability and validity measures. As the authors point out it is economical and efficient since it is a group test. Such a cloze test certainly would provide interesting data for a multitude of research studies into how children process various types of written language. The whole notion of using cloze data in a qualitative manner could provide data for studies of text cohesion, children reading as authors, etc.

However, if, as DeSanti hopes, this cloze test will be used for "diagnostic" purposes - to determine children's reading strategies as a basis for designing appropriate intervention i.e. determining children's instructional needs, there is at least one major flow as I see it.

Cloze substitutions constitute an artifical "error" if you will. To my knowledge there is no established predictive positive relationship between a child's patterns of miscues and his/her cloze substitution/insertions. In other words, for purposes of determining instructional needs, I'm not sure what you have after you have analysized a child's insertions on such a reading task.

While the research uses of such an approach with cloze seems quite promising, I would caution against its use as an indicator of instructional need(s), at least until further research has been completed.

CONTENT AREA READING: PREACHING AND PRACTICING

RICHARD TELFER
University of Wisconsin-Whitewater

Content reading instruction is included as part of teacher preparation, with the expectation that teaching will change because of the instruction. Yet, Ratekin, Simpson, Alvermann, and Dishner (1985) reported that "recommendations made by textbook authors were rarely implemented" (p. 436) in classrooms. Other studies (Gehrke, Schaefer, & Schlick, 1982; Smith & Feathers, 1983) have indicated similarly. According to these descriptions, even if knowledge and attitudes about content reading strategies are changed through coursework, teachers' use of strategies may not be changed. Specifically, Ratekin et. al. concluded that teachers in general do not use content reading strategies. In their explanations of this condition, they suggested that teachers adjust the content reading model to fit their situations, find content reading strategies to be of little value, or fail to use the text as the primary source of concept development.

However, the explanations offered by Ratekin et. al. may not be the only plausible ones. One alternative explanation is that the observed lack of use of content reading strategies may be related to lack of teacher knowledge of those strategies. Teachers who do not know about content reading strategies logically would be unlikely to use those strategies. And teachers who have not taken content reading classes or who have not taken the courses recently would be lacking knowledge of at least some strategies. Another possibility is that content reading instruction has an effect on teaching strategies, but this effect is more general modification of the activities used than the adoption of a set of "content reading strategies." That is, teachers who have taken content reading courses may use different types of activities in general than those who have not taken content reading courses.

This study is designed to examine these possibilities. First, is there a relationship between content reading coursework and knowledge of content reading terminology? Specifically, do teachers who have taken content reading courses differ from teachers who have not taken content reading courses in amount of knowledge about content reading? In addition, is any difference in knowledge related to time? Do teachers who have taken content reading coursework have a different level of content reading knowledge from those who have not even if the coursework was taken several years previously. It would be expected that teachers having taken content reading courses would score higher on a test of knowledge of content reading terminology than teachers who have not. It would also be expected that teachers having taken content reading coursework more recently would score higher.

Second, are the teaching behaviors of teachers who've taken content reading coursework different from the teaching behaviors of teachers who have not taken content reading coursework? This question differs from that addressed by Ratekin et. al. in one fundamental way. While Ratekin et. al. looked at the instructional strategies used by teachers and concluded that teachers did not use the strategies recommended in content reading textbooks, they did not discuss whether or not these teachers had taken content reading coursework.

If a comparison of teachers who have and have not taken content reading classes is made, activities that seem suited to the three main content reading assumptions identified by Ratekin et. al. may be expected to be reported more frequently and for longer periods of time by teachers who have taken content reading classes. That is, activities that indicate variation of organizational settings and instructional methods; activities that focus on students' readiness for learning, acquisition of information, and internalization of concepts; and activities that involve a variety of instructional resources would be seen more often and for longer periods of time among teachers who have taken content reading classes. In this study, therefore, group work and discussion, introduction of material, discussion of vocabulary, and summarizing material would be expected to be reported more frequently and for longer periods of time by teachers who have taken content reading classes than by teachers who have not.

METHOD

Subjects

The subjects were middle school and high school social studies teachers from throughout the state of Kansas. They were randomly selected using a state of Kansas educational directory and a list of random numbers. Since no comprehensive listing of social studies teachers in the state of Kansas was available, letters were sent to the principals at the selected high schools and middle schools. These letters asked the principals to direct a preliminary survey to social studies teachers at their schools. Based on responses to this preliminary survey, twenty-four teachers were chosen for further participation. Twelve participants were randomly chosen from those teachers who expressed willingness to participate and who had taken a content reading course. The other twelve participants were randomly chosen from those teachers who expressed willingness to participate and who had not taken a content reading course.

Instruments

The preliminary survey asked for several types of information. It asked for subject specialty, grade level(s), and years of teaching experience. It also asked whether the teacher had taken a content reading course and whether the teacher had taken any other reading methods course. For each of the courses the date when the course was taken was also requested. In addition, respondents were also asked whether they would be willing to participate further, by taking a test and/or by keeping a record of instructional activities.

A 25-item-multiple-choice test of knowledge of content area reading terminology was developed by the examiner. The development of the test consisted of compiling a list of terms associated with content area reading, verifying this list, and selecting and developing items. The list of terms was developed by looking through the tables of contents and indexes of five content area reading textbooks. Terms were selected if they were thought to represent concepts that could be manifested in instructional activities and if they were seen as associated primarily with content area reading rather than some other area of reading. Verification of this list consisted of checking to see if the terms were used in the majority of the textbooks surveyed. Items were selected and developed according to the following pattern. Items were selected from a previous test of teachers' knowledge (Telfer, 1984) if the items could be described as focusing on content reading terminology. Additional items were also developed.

A record form was developed for the teachers to use when recording instructional activities. Since knowledge of the specific activities mentioned or suggested in the test was not assumed, teachers were asked to record the general types of activities they used. In order to make this recording process easier to complete and analyze, two sets of activities were listed as common activities, those describing general types of behavior (such as group work or lecture) and those describing more specific behaviors (such as teacher explanation or student reading). Teachers were asked to record the instructional activities they used over a period of five consecutive days. Teachers were also told that they need not limit themselves to the activities listed on the form.

Procedures

Middle school and high school social studies teachers from throughout the state of Kansas were asked to respond to a survey that asked for basic information about subjects taught, grade levels taught, content reading courses taken, other reading courses taken, and years of teaching experience. The teachers were also asked whether they would participate further in the study.

Teachers who expressed their willingness to participate further were divided into two groups: those who had taken a content area reading course and those who had not. From these two groups a total of twenty-four teachers were chosen to participate further, twelve from each group. These teachers were mailed a test of content area reading terminology and an instructional activities record sheet, along with instructions for both.

The scores on the test for the two groups were compared. In addition, correlations between test scores and years of teaching experience were figured for both groups.

The records of instructional activities were examined using the following procedure. First, a percentage of time was figured for each type of instructional activity for each teacher. Then these percentages of time were graphed to give a profile for each teacher. Once the percentages of time were figured for all of the teachers, the totals were analyzed in two ways. Percentages of time were treated as standard scores and averaged across each group. From this a profile of the group was developed. Total time spent by each group for each activity was also calculated and graphed.

RESULTS

The following information was gathered from responses to the preliminary survey. Surveys were mailed to 200 schools. A total of 131 surveys were returned, with 72 respondents (55%) expressing willingness to participate further. At the same time, 35 (27%) indicated that they had taken a course in the teaching of reading in content areas, including 19 who expressed willingness to participate further. In addition, 31 respondents (24%) indicated that they had taken one or more other courses in the teaching of reading and 8 who would participate but who had not taken content area reading).

Respondents reported taking content area reading courses between 1969 and 1985. The median year when courses were taken was 1980. The range of years was the same for those willing and unwilling to participate. The medians were different, however. For the group of teachers willing to participate further, the median year when a content reading class was taken was 1980. For the other group, the median year was 1978.

The number of years of teaching experience completed ranged from no experience to 35 years. The mean number of years teaching experience completed was 12.79. The groups of teachers differed, though. The mean number of years of teaching experience for those teachers willing to participate was 11.16, while the mean years of teaching experience for the other group was 15.21. An additional difference was also noted. The mean years teaching experience for those who have taken content reading coursework was 10.80, while the mean for those who have not was 13.71.

The teachers who indicated that they would participate were divided into two groups, one consisting of those who had and the other consisting of those who had not taken a content reading course. From each of these groups 12 individuals were randomly chosen to take the test and keep a record of instructional activities. An examination of the characteristics of the 12 chosen in each group indicated that the characteristics of the smaller (12 person) groups were roughly equivalent to the larger groups.

Mean scores on the 25-item test of knowledge of content reading terminology were 16.0 for those who had taken content reading courses (YES-CAR) and 14.9 for those who had not (NO-CAR). The difference is not significant (t[14]=0.71).

Correlations between content area reading terminology scores and years of teaching experience were -0.53 for the YES-CAR group and -0.26 for the NO-CAR group. The correlation between test scores and years since taking a content reading course was -0.48 for the YES-CAR group.

Visual inspection of the average percentages of time spent in various activities (See Table 1) shows differences of 5% or more in five of the categories. The NO-CAR group had higher percentages in student reading, individual work, and lecture. The YES-CAR group had higher percentages in small group work and discussion. A similar pattern is seen when total time is graphed, although the largest differences are in student reading and discussion.

DISCUSSION

Responses to the initial survey suggest that the participants in the study may be representative of social studies teachers in the state of Kansas. The teachers who expressed willingness to participate were similar to the teachers in the total sample. Similarities were seen in courses taught, grade levels taught, percentage who had taken a content reading course, and percentage who had taken another reading course. The major area of difference was in years of teaching experience. Those who expressed willingness to participate had taught on the average four fewer years than those who declined participation. While this difference must be noted, the group of teachers who expressed willingness to participate were clearly experienced teachers.

The average number of years teaching experience also seems to be a difference between those who have and those who have not taken content reading courses. Those who have taken content reading courses have almost three fewer years teaching experience than those who have not taken the courses. While this may be expected, with the relative recency of emphasis on content reading courses, the difference in teaching experience remains a confounding factor.

TABLE 1

AVERAGE PERCENTAGE OF TIME IN VARIOUS CLASSROOM ACTIVITIES

Activity	YES-CAR	NO-CAR
Teacher Explanation	9%	10%
Introduction	5%	8%
Vocabulary Discussion	2%	4%
Correcting Answers	4%	4%
Summarizing	6%	3%
Reading	12%	19%
Answering Questions	6%	7%
Problems/Worksheets	12%	8%
Individual Work	18%	24%
Small Group Work	16%	6%
Lecture	6%	14%
Testing	6%	5%
Discussion	21%	15%
Other	8%	11%

Percentages do not total to 100% because of overlapping of categories.

Scores on the test of content reading terminology may show two patterns. The small advantage for the YES-CAR group on the test of knowledge of content reading terminology was in the direction expected, but of less magnitude than expected. At the same time, there was a rather strong negative correlation between teaching experience and test score. This negative correlation was strongest for the YES-CAR group, suggesting that what was learned was then forgotten. An explanation of the lack of use of content reading strategies may be that teachers do not currently know the strategies. And, if they do not know the strategies, they cannot use them.

The reporting of activities used indicates some preliminary patterns. Some of the patterns were as expected: The YES-CAR group reported more group work and more discussion. Other patterns were different from those expected: The NO-CAR group spent more time introducing vocabulary; neither group spent much time discussing vocabulary. If content reading instruction is having an influence, it may be by encouraging teachers to shift from one type of instruction to another. If so, greater amounts of group work and discussion may be manifestations of content reading instruction.

The conclusions in this paper are necessarily tentative, but they can indicate directions for further research. A second phase of this study will look at the actual practices in the classroom of selected teachers who participated in the first portion of the study. Observations will look first at the actual use of activities according to the categories used in the self-report. Further analysis will look at the specific types of activities suggested in the test of content reading terminology.

REFERENCES

Gehrke, N.; Schaefer, J.; & Schlick, K. (1982). Reading in content areas: A comparative study of attitudes, perceptions, practices of preservice and inservice secondary teachers. (ERIC Document Reproduction Service No. ED 215 331).

Ratekin, N.; Simpson, M.; Alvermann, D.; & Dishner, E. (1985). Why teachers resist content reading instruction. Journal of Reading, 28, 432-437.

Smith, F. & Feathers, K. (1983). Teacher and student perceptions of content area reading. Journal of Reading, 26, 248-254.

Telfer, R. (1984). The development and initial validation of a test of teachers' knowledge of the teaching of reading comprehension. (Unpublished Doctoral Dissertation, University of Wisconsin-Madison.

CONTENT READING TEXTS USED IN DEVELOPMENT OF TEST

Estes, T. & Vaughan, J. Jr. (1985). Reading and learning in content classroom (2nd ed.) Boston: Allyn and Bacon.

Karlin, R. (1984). Teaching reading in high school (4th ed.) New York: Harper.

Roe, B.; Stoodt, B.; & Burns, P. (1983). Secondary school reading (2nd ed.). Boston: Houghton Mifflin.

Singer, H. & Donlan, D. (1980). Reading and learning from text. Boston: Little, Brown.

Vacca, R. (1981). Content area reading. Boston: Little, Brown.

PRELIMINARY DEVELOPMENT OF A SCREENING INSTRUMENT FOR LEARNING DISABILITIES GRADES 2-8

THOMAS CLOER, JR.
Furman University

An increasing number of professionals in the field of reading have expressed distress regarding use of terms such as "learning disability," "dyslexia," "neurological dysfunction," and "developmental lag." These terms seem to be generally used to refer to students that are not doing as well a someone thinks they should.

Groups of students fail to achieve in reading because of a plethora of reasons. These reasons include physical, social, emotional, experiential, motivational, and intellectual variables. The difficulty in discussing dyslexia or learning disability is exacerbated by differences in interpreting the terms. A learning disability in this paper will refer to a visual or auditory perceptual deficit as measured by the short-term ability to remember and write nonsense syllables presented visually, and the ability to hear speech sounds in spoken words.

Since Public Law 94-142 was signed into effect in 1978, the area of learning disabilities has taken on added significance. Schools are now compelled by law to offer specialized instruction to all students identified as having learning disabilities. A major part of this new emphasis involves the identification and documentation of bona fide learning disabled students. This has been a major problem since virtually no instruments are available for teachers to use that furnish valid and reliable data deemed useful for delineating disabilities. Most tests are clinically oriented and are not well supported by empirical data that show a relationship between the constructs measured and school tasks. The goal of the research discussed here is to develop an instrument supported by basic research that will efficiently and effectively identify perceptually disabled students in relation to reading.

Barrett (1965), Vernon (1973), and Durrell (1963) have documented the need for accurate visual perception of letters and words in reading achievement. Karlin (1980) concluded that the ability to discriminate visually among letter and word forms has a greater influence on recognizing words than the ability to discriminate among non-verbal forms. Dykstra (1966), Durrell and Murphy (1953), and Morency (1968) have also demonstrated the high correlation between auditory perceptual abilities and reading achievement. Durrell and Murphy (1953) have maintained that the auditory perception of sounds in spoken words is crucial to early reading achievement.

Cloer (1979) attempted to analyze seven different vocabulary lists to ascertain which spellings of English phonemes were most frequent. The study was based on the assumption that by determining the frequency of occurrence of various spellings for different phonemes, it would be possible to determine which sound-symbol relationships would most frequently help students in decoding words unfamiliar to them. Cloer (1983) also attempted to identify and analyze the 500 most frequently occurring multisyllabic words in order to determine which visual patterns appeared most frequently and which visual syllabication rules could veritably assist with pronunciation. The findings from these studies were used in constructing the instrument.

Method and Subjects

The short screening test was designed to include two subtests, visual and auditory. The visual test consisted of 15 items for grades 1, 2, and 3. Grades 4, 5, 6, 7, and 8 were given 20 items. The visual test asked students to study nonsense words of increasing length for 5 seconds, and then write from memory what they had seen. The nonsense words were similar to English words, and contained the visual patterns that occurred most frequently according to the previous studies. The nonsense words were presented on standard typing paper, 8 1/2 x 11 inches. The tallest letters were 2 inches high.

The auditory subtest consisted of 20 items, and required students to listen and watch on their individual test sheets as the examiner pronounced an English word and four options with three distractors. The student was to underline the word that contained the same vowel sound as the stimulus word. Most distractors were visually irregular in order to measure the students' use of auditory perception as revealed through identifying vowel sounds that are the same. The student could not determine the vowel sounds, and successfully match them, using the visual patterns alone. This test was untimed, but pace was controlled by the examiner who read each item twice. The two subtests, visual and auditory, were combined to give a total score. (Copies of the complete tests are included at the end of this article.)

A total of 503 students in grades 2 through 8 were administered tests. There were 76 students in grade 2, 60 in grade 3, 70 in grade 4, 55 in grade 5, 70 in grade 6, 88 in grade 7, and 84 in grade 8. The screening tests were completed by pupils in a total of 21 classes in two elementary schools and a middle school who participated in the pilot study. The teachers administering the tests were provided training in a short session.

Several standardized reading tests were used to determine concurrent validity. These included: Gates-MacGinitie Reading Test (Total Score) for grade 2, Stanford Achievement Test (Reading Words) for grade 3, South Carolina Test of Basic Skills (Reading Total) for grades 3 and 4, Comprehensive Test of Basic Skills (Reading Total) for grade 5, and Botel Word Opposites for grades 6, 7, and 8.

Results

Validity

Table 1 presents the grades tested, means, and standard deviations for scores on the screening test.

Table 1

Means and Standard Deviations for Screening Test, Grades 2-8

	Mean			Standard Deviation		
Grade	Vis.	Aud.	Total	Vis.	Aud.	Total
2 (N = 76)	10.27	9.39	19.67	2.68	5.19	7.01
3 (N = 60)	10.15	12.63	22.78	2.95	5.11	7.16
4 (N = 70)	11.95	11.42	23.54	3.64	5.40	7.88
5 (N = 55)	12.93	15.67	28.24	5.09	4.22	7.14
6 (N = 70)	15.48	14.60	30.80	3.37	5.12	7.39
7 (N = 88)	16.35	15.06	31.42	3.35	4.08	6.83
8 (N = 84)	16.50	15.14	31.71	3.21	4.60	6.57

Table 2 provides the correlations between the screening test and the standardized reading tests used in each grade.

Table 2

Correlation between Screening Test and Standardized Reading Tests

		Pearson Product Moment Correlation		
Grade	Test Administered	Vis.	Aud.	Total
2	Gates-MacGinitie	.66*	.76*	.79*
3	Stanford Achievement	.64*	.65*	.75*
4	South Carolina Test of Basic Skills	.46*	.63*	.64*
5	Comprehensive Test of of Basic Skills	.27**	.58*	.58*
6	Botel Word Opposites	.60*	.60*	.69*
7	Botel Word Opposites	.46*	.70*	.72*
8	Botel Word Opposites	.32**	.42*	.45*

*$p<.001$.   **$p = .02$.

The reader will note that in every grade there was a significant correlation between each screening subtest and the standardized reading tests used as criteria for concurrent validity. There was a low correlation between the visual subtest of fifth graders and the Comprehensive Test of Basic Skills, $r(53) = .27$, $p = .02$. Another low correlation existed between the visual subtest for eighth graders and the Botel Word Opposites Test, $r(82) = .32$, $p = .02$. All Total scores from the screening test were significantly correlated to the different reading tests used as the criteria at or beyond the .001 level of confidence. The screening test accounted for as much as 62% of the variance in the second grade and as low as 20% of the variance in the eighth grade. This is not surprising since perception should be secondary to higher cognitive processes in accounting for the variance at this level. Surprisingly, at the sixth and seventh grade levels, the total screening test accounted for 47% and 51.8% of the variance respectively. This may have occurred because the Botel Word Opposites Test is really a test of word recognition, and word recognition was a problem for many of these students.

Reliability

Spearman-Brown reliability coefficients were obtained for each grade level. The standard procedure for calculating split-half reliability, correlation of raw scores on odd versus even items, was used. The resulting split-half correlations were corrected for half-test length by the Spearman-Brown formula. Table 3 gives the Spearman-Brown coefficients of internal consistency. These coefficients reflect measurement error resulting from characteristics of the particular sample of items used in the test.

Table 3

Split-Half Reliability Coefficients for Visual and Auditory Subtests and Total Scores, Grades 2-8

Grade	Visual	Auditory	Total
2	.73 (N = 149)	.88 (N = 163)	.88 (N = 119)
3	.85 (N = 74)	.87 (N = 73)	.88 (N = 71)
4	.81 (N = 108)	.91 (N = 113)	.91 (N = 102)
5	.75 (N = 39)	.83 (N = 39)	.85 (N = 38)
6	.89 (N = 91)	.88 (N = 83)	.88 (N = 42)
7	.78 (N = 29)	.88 (N = 29)	.88 (N = 29)
8	.85 (N = 119)	.85 (N = 115)	.87 (N = 89)

It can be seen from these data that the subtest coefficients ranged from .73 on the visual subtest for second graders to .91 on the auditory subtest for fourth graders. Reliability coefficients for the total test, visual and auditory combined, were in the high eighties to the low nineties.

Discussion

With the results summarized according to the different grades, the results indicate that the items measured do have concurrent validity in relation to standardized reading tests being presently used. The reliability data indicate that measurement error resulting from characteristics of the particular sample of items is not a significant problem with this screening instrument.

It is somewhat surprising that this much of the variance in reading achievement is accounted for by a brief screening of perception. There is such a plethora of experiential, social, emotional, physical, intellectual, and motivational variables included in this variance. The short screening instrument shows promise of identifying students in grades two through eight that might have severe perceptual problems. The instrument might also help with the current imbroglio that often has two people at convention such as the American Reading Forum debating vehemently and intransigently only to discover that they were talking about different things, and were majestically arguing for the same point. There are students who have difficulty learning how to read even though social, emotional, physical, or motivational problems appear to be minimal. There are students who do not hear sounds in spoken words as well as their peers. There are students who have severe difficulty remembering the letter sequences that constitute words. These subjects are the ones that this instrument might help identify, so that special assistance can be given. The writer hopes that colleagues in the Forum will do further research with this instrument. Performance by good and poor readers could be examined. The instrument might be employed in isolation or as part of a battery to identify learning disabled and provide diagnostic data to aid in remediation.

SCREENING INSTRUMENT FOR LEARNING DISABILITIES

Visual Subtest-(Read the Directions for the students.)

Directions: I will show you cards on which nonsense words are printed. Study each word until I remove the card. Then write as much of the nonsense word as you remember. Each card will be shown for five (5) seconds. Please do not begin to write until the cards are removed. The cards cannot be shown a second time.

1. hib
2. yad
3. feg
4. drub
5. rame
6. sarn
7. blesy
8. folet
9. shaty
10. grelet
11. scofle
12. liffeg
13. pafhups
14. smomter
15. dinonast

*16. thiolary
*17. mibedaffe
*18. chiffopof
*19. tenisilext
*20. pimosuffes

*Administer only the first fifteen items for students in Primary grades (1-3). Administer all twenty items above grade 3.

Auditory Subtest-(Teacher reads aloud as student follows.)

1. The vowel sound heard in $\boxed{\text{sn}\underline{\text{i}}\text{ff}}$ is the same vowel sound heard in:

   a. fir b. view c. light d. give

2. The vowel sound heard in $\boxed{\text{p}\underline{\text{i}}\text{ck}}$ is the same vowel sound heard in:

   a. thief b. ridge c. first d. wife

3. The underlined vowel sound heard in the first part of $\boxed{\text{sn}\underline{\text{i}}\text{ck er}}$

is the same vowel sound heard in:

a. high b. build 3. bike 4. suite

4. The vowel sound heard in speck is the same vowel sound heard in:

a. key b. bead c. head 4. here

5. The vowel sound heard in get is the same vowel sound heard in:

a. thread b. seal c. earth d. heed

6. The vowel sound heard in at is the same vowel sound heard in:

a. Claus b. laugh c. mate d. fall

7. The vowel sound heard in tack is the same vowel sound heard in:

a. shall b. base c. calm d. heart

8. The vowel sound heard in mud is the same vowel sound heard in:

a. full b. rule c. use d. blood

9. The underlined vowel sound heard in the first part of ug ly is the same vowel sound heard in:

a. rough b. should c. due d. flu

10. The underlined vowel sound heard in the first part of hot ter is the same vowel sound heard in:

a. booth b. owl c. wad d. look

11. The vowel sound heard in her is the same vowel sound heard in:

a. there b. earth c. here d. thread

12. The underlined vowel sound heard in the first part of ho tel is the same vowel sound heard in:

a. cloth b. out c. of d. old

13. The underlined vowel sound heard in the first part of fe male is the same vowel sound heard in:

a. end b. piece c. shelf d. thread

14. The underlined vowel sound heard in the last part of par ty is the same sound as the "y" in:

a. any b. yellow c. gym d. cry

15. The vowel sound heard in snake is the same vowel sound heard in:

a. have b. chap c. chalk d. sleigh

16. The vowel sound heard in Ford is the same vowel sound heard in:

a. rock b. world c. floor d. words

17. The vowel sound heard in beet is the same vowel sound heard in:

a. been b. key c. her d. cent

18. The vowel sound heard in bean is the same vowel sound heard in:

a. bread b. check c. break d. tea

19. The vowel sound heard in bark is the same sound heard in:

a. care b. are c. stare d. raft

20. The underlined vowel sound heard in the first part of glit ter ing is the same vowel sound heard in:

a. thief b. shirt c. hike d. been

References

Barrett, T. C. (1965). The relationships between measures of pre-reading visual discrimination and first grade reading achievements: A review of the literature. Reading Research Quarterly, 1, 51-76.

Basic Skills Assessment Program--Reading. (1981). S.C. State Department of Education. Columbia: Educational Assessment Office.

Botel, M. (1978). Botel word opposites test. Chicago: Follett.

Cloer, C. T. (1979, Spring). A phonics frequency count. Search-Newsletter of Research Committee for S.C. Council, I.R.A., pp. 1-7.

Cloer, C. T. (1983). An analysis of the American Heritage word list--Implications for teaching beginning reading. In G. McNinch (Ed.), Reading research to reading practice: Third yearbook of the American Reading Forum (pp. 47-48). Athens: American Reading Forum.

Comprehensive tests of basic skills. (1974). Monterey, CA: McGraw-Hill.

Durrell, D. D., & Murphy, H. A. (1953). The auditory discrimination factor in reading readiness and reading disability. Education, 73, 556-560.

Durrell, D. D., & Murphy, H. A. (1963). Reading readiness. Journal of Education, 146, 3-10.

Dykstra, R. (1966). Auditory discrimination abilities and beginning reading achievement. Reading Research Quarterly, 1, 5-34.

Karlin, R. (1980). Teaching elementary reading (3rd ed.). New York: Harcourt Brace Jovanovich.

MacGinitie, W. (1978). Gates-MacGinitie reading tests. Burlington, MA: Riverside Publishing Co.

Morency, A. (1968). Auditory modality, research, and practice. In H. K. Smith (Ed.), Perception and reading (Vol. 12). Newark, DE: International Reading Association.

Stanford achievement test. (1973). New York: Psychological Corporation.

Vernon, M. D. (1973). Visual perception and its relationship to reading: An annotated bibliography. Newark, DE: International Reading Association.

EFFECTIVENESS OF INFORMAL TEACHER OBSERVATIONS IN DETECTING STUDENT VISUAL DIFFICULTIES

JOHN ALVAREZ
Edison Middle School

SHARON KOSSACK
Florida International University

Reading is a visual act adversely affected by hyperopia (far sightedness), binocular incoordination, and fusion (Bond and Tinker, 1973). In particular, hyperopia (near point difficulty) is found in poor readers (Hartlage, 1976). It would seem helpful to remove the obstacle of visual difficulty to enable disabled readers to function with comfort in a reading situation, since reading itself presents enough frustration. Earlier researchers (Robinson, 1968) concluded incidences of visual defects among reading disabled was not disproportionate. This suggested that vision difficulties do not significantly contribute to reading failure; however some researchers have noted evidence to the contrary. Bond and Tinker (1973) assert that binocular incoordination, fusion, and farsightedness (hyperopia) seem to affect the reading act. Hartlage (1976) found a greater incidence of hyperopia among poor readers. Since there are differences among researchers, it would appear a system of classroom screening, referral, and appropriate prescription could be processed for identifying visual defects of students for every school. But, are they practical or effective? Are they available for usage?

For instance, formal screening devices such as the Keystone Telebinocular might detect vision problems; but are not economically available for classroom use.

Another alternative, such as planned public school screenings by health personnel are ineffective. Ekwall (1983) indicated 50 percent of the disabled readers assessed in the University of Texas Reading Clinic had visual problems that were undetected through normal school screening procedures. Many schools only apply the Snellen, which is an instrument designed to detect farpoint vision problems of students. The Snellen does not analyze nearpoint reading tasks.

Another alternative is teacher observation checklists (Wilson, 1977). These checklists present an economical alternative, but are they a viable, effective, and valid means of referring students? Also, do checklists detect actual visual deficiencies.

This study was conducted to investigate the usage of a checklist for teachers in detecting visual problems of students. A copy of the checklist employed is included at the end of this article.

Method

The study was divided into three parts. They were:

1. Teachers were trained to use a checklist to observe indicators of visual difficulties.

   Teachers were asked to observe their students and identify anyone that showed evidence of visual difficulty, using a checklist as a guide. The checklist used was the St. Louis Optometric Association's. (Wilson, 1981)

2. The screening results of randomly selected groups of low and average achieving students were compared.

   The screening results of the low and average achieving students were compared by using Chi Square. This validated the incidence of hyperopia.

3. The prescription effectiveness was evaluated.

TABLE I: Assessment of Higher Incidence of Visual Difficulties in Lowest Achieving Subgroup

	Satisfactory	Unsatisfactory
Observed	29	11
Expected	28	12

N = 40 df = 1

Chi Square = .119 $p < .01$

The identified students were then screened with a telebinocular and the results were verified by an opthamologist in order to judge the relative accuracy of the teacher observations.

TABLE II: Incidence of Hyperopia (Near point acuity)

	Satisfactory	Unsatisfactory
Observed	67	27
Expected	89	5

N = 94 df = 1
Chi Square = 102.23 $p > .01$

Glasses were provided at no charge to students who needed them. Then as a means of evaluating the effectiveness of the prescriptive plan, students were observed to see if they wore their glasses on a consistent basis.

Data Source

Inner city low-socio economic level seventh graders from a multi-ethnic (Black:Hispanic:White ratio of 80:15:3) school. Approximately 90% of the student body had been described as reading two or more years below grade level.

Results

Effectiveness of Teacher Observation

Of the students referred by teachers on the basis of the checklist, 90% were diagnosed as needing glasses, as verified by medical doctors. This suggests that trained teachers can act as effective visual screening agents. Teacher observation referrals did identify students who needed visual adjustment. Further study is needed to determine if the visual screening identified *all* of the students needing glasses; no effort was made to determine if students were underreferred.

Incidence of Specific Visual Difficulties

Telebinocular results of low readers (Chapter I subpopulation; i.e. students scoring in stanines 1 and 2 on the *Stanford Achievement Test*) were compared with the general school population in terms of detected visual deficiencies. There were no significant differences in the percentage of students with unsatisfactory overall vision (Chi Square .119; df = 1, nonsignificant at .01 level), confirming the Robinson (1968) study.

Closer analysis, however, showed significantly higher numbers of students in the general school population tested who showed hyperopia (nearpoint acuity difficulties). Results from the 1969 Coleman study were used to determine significance levels: data had suggested that myopia (nearsightedness) is 5 times as prevalalent as hyperopia (farsightedness) for fifth graders and increases rapidly after age 5. Hence, the ratio 5:1 was used as a minimally acceptable figure for these seventh graders (chronological age 12) which incidence of hyperopia exceeded expected levels (Chi Square of 102.23; df = 1; .01 level of significance) reported in Coleman (1969) study. These results were consistent with Hartlage's (1976) findings of significant levels of hyperopia among poor readers.

Effectiveness of Prescription

Teachers were asked to observe students who received unsatisfactory scores on the visual assessment to determine how many followed through wearing glasses (prescription). It is important to note that glasses were provided at no charge whenever necessary. As a means of evaluating the actual effectiveness of the school-wide visual referral process, teachers observed those who did secrue glasses to determine how many actually wore them in class.

Only 25% of the students needing glasses actually secured them, in spite of free prescriptions. Of the group who obtained glasses, only 4% wore them in classes. Students needing glasses, who read more than three grade levels below placement, were especially observed. NONE of this smaller group wore glasses in class. Visually deficient students who had glasses were not wearing them during classroom work. Check-up and reinforcement procedures are needed to assure the effectiveness of visual prescription of wearing glasses.

Conclusions

Results were consistent with earlier findings. The low achieving students do not have a higher incidence of visual difficulties than the standard population; however, incidence of hyperopia was significantly higher for this inner city junior high school population than would be expected in general.

From these results, it would appear that a trained and observant classroom teacher who uses a visual checklist can detect significant visual needs. Ninety percent of teacher observation generated referrals were shown to require prescriptions for visual difficulties. A one-month-a-year emphasis on observing children for visual difficulties at all levels would likely detect visual difficulties early enough that they would not pose the problems that have persisted, undetected, for high school students.

A greater problem involved the reluctance of the students to wear corrective lenses when prescribed. A system of reminder because the glasses-wearing habit is a new one and must be learned) and motivation (to offset the stigma and peer pressure against wearing glasses) needs to be revised to assure that students needing glasses will, in fact, make use of them.

Recommendations

Because teacher observation was shown to be effective in identifying a large percentage of students with visual problems, teachers should be trained to observe symptoms of visual problems early in the year. Early in the year an observation checklists and referrals in each school for one month would likely identify visual difficulties early. If students are identified and aided, it could influence their reading abilities.

Over fifty junior high (seventh grade) level students were found to need glasses following an assessment made after six years of schooling without glasses. The only means of detecting visual problems with these students was an occasional Snellen screening. This testing apparently was administered to a few of the students. The Snellen measures only farpoint (ability to see blackboard work, bulletin boards, etc.) and is not intended to detect students needing adjustment for nearpoint (close vision, as in reading a book) difficulties. Because telebinocular screenings are relatively easy to give and most local Health Departments will loan out the equipment, schools should have a standard and systematic means of early detection and referral which should include teacher-observation referral and follow-up telebinocular screening. It appears that such a screening should not be limited to the poor-achieving subgroups, but should be conducted on a school-wide basis, since the incidence of visual difficulty was as high in the standard school population as it was in the Chapter I group.

BIBLIOGRAPHY

Bond, G. and Tinker, M. (1973). Reading Difficulties: Their Diagnosis and Correction (3rd. Ed.). Appleton-Century-Crofts.

Cienkus, Robert. (1974). Selected Visual Abilities As Correlates to Reading Achievement. (Doctoral Dissertation, Loyola University of Chicago). Dissertation Abstracts International 35, 2080.

Coleman, Howard M. (1968). Visual Perception and Reading Dysfunction, Journal of Learning Disabilities, I, 116-123.

Drebry, Catherine. (1979). Vision Problems and Reading Disability, Reading Teacher, 32, 787-795.

Ekwall, E. and Shanker, J. (1983). Diagnosis and Remediation of the Disabled Reader (2nd Ed.). Boston: Allyn and Bacon.

Hartlage, Lawrence. (1976). Vision Deficits and Reading Impairment. In Basic Visual Processes and Learning Disability, G. Leisman (Ed.), Springfield, IL: Thomas Publisher, pp. 151-161.

Martin, Harold. (1971). Vision and Its Role in Reading Disability and Dyslexia, Journal of School Health, 41: 468-72.

Ritty, Michael. (1979). Assessing and Alleviating Visual Problems in Schools, Reading Teacher, 32, pp. 796-802.

Robinson, Helen. (1946). Visual Efficiency and Reading Status in the Elementary School, Clinical Studies in Reading, III, Chicago: University of Chicago Press, p. 49-65.

Wilson, Robert M. (1981). Diagnostic and Remedial Reading for Classroom and Clinic, Charles E. Merrill Books, Inc., Columbus, Ohio, p. 74.

OBSERVATION CHECKLIST OF VISUAL PROBLEMS

DIRECTIONS: Please use this checklist to observe your students while they are working in your classes. Refer for follow-up visual screening any student that shows four or more of the following symptoms. Refer any student who shows * (starred) symptoms.

____ *Losing place while reading.
Slow reading or word calling.
Poor perceptual ability, such as confusing o and a, n and m.

____ *Avoiding close work.

____ *Poor sitting posture and position while reading.
Fatigue or listlessness after close work.
Inattentiveness, temper tantrums, or frequent crying.
Complaint of blur when looking up from close work.
Seeing objects double.
Headaches, dizziness, or nausea associated with the use of the eyes.

____ *Body rididity while looking at distant objects.
Undue sensitivity to light.
Crossed eyes--turning in or out.
Red-rimmed, crusted or swollen lids.
Vocalizing during silent reading, noticed by watching lips or throat.
Reversals persisting in grade 2 or beyond.
Inability to remember what has been read.
Complaint of letters and lines "running together" or of words "jumping."

____ *Holding reading closer than normal.

____ Frowning, excessive blinking, scowling, squinting, or other facial distortions while reading.

____ Excessive head movements while reading.
*Writing with face too close to work.
Frequent sties.
Watering or bloodshot eyes.
Burning or itching eyes or eyelids.

____ *Tilting head to one side.

____ *Tending to rub eyes.
Closing or covering one eye.
Frequent tripping or stumbling.
Poor hand and eye coordination as manifested in poor baseball playing, catching and batting, or similar activities.

____ *Thrusting head forward.

____ *Tension during close work.

THE MANDATE: EVALUATION OF TEACHER CANDIDATES BY USE OF COMPETENCY TESTS

RONA F. FLIPPO
Fitchburg State College

The era of teacher competency testing is here. The evidence is all around: in the professional journals, newspapers, and newsletters; in the media; and in the public schools. The evaluation theme of this reading conference, with a strand focused on teacher assessment, is evidence that many reading educators have acknowledged the magnitude of this era.

The mandate to evaluate teacher competencies has become a full-scale epidemic, much like the widespread minimum competency testing of children in the 1970's. In the 70's school-aged youngsters and their teachers were penalized if the youngsters did not pass the minimum competency tests forced on them. Now, in the 80's, entry level teachers and their colleges are penalized if the entry level teachers do not pass the minimum competency forced on them. As in the past, minimum competencies are the focus, even though evidence indicated by the recent major studies in education show that schools and students are still not up to par. As a matter of fact, schools and students may be weak because of the mediocrity that the minimum competency testing tends to promote. Research has indicated that this type of widespread testing can often retard rather than advance the interests of students, since bureaucratically sanctioned testing tends to place more focus on the tests than on the substance of what the tests are intended to measure (Haney, 1984). However, the education decision makers have not learned; they continue to promote minimum competency testing.

The widespread breadth and scope of the teacher competency testing era has been documented in many papers (Flippo, 1985; Flippo & Foster, 1984; McCarthy, 1985; Sandefur, 1985; Schnittjer & Flippo, 1984), and in other recent publications (e.g., Teacher Competence, 1985, one of the books in Phi Delta Kappa's Hot Topic series). This testing is either taking place, being developed, being planned, or being discussed in almost every state in the United States. Bosworth (1985) noted that only six states did not have some sort of teacher testing implementation, development, or discussion taking place. Those states were Alaska, North Dakota, Rhode Island, Utah, Vermont, and Wyoming.

The Effect on the Field of Reading Education

Those involved with reading education have for sometime looked beyond the minimum skill abilities of students and asked for an emphasis on a more holistic evaluation of students. Reading educators have argued against testing minimum skill or subskill proficiency and instead have argued for evaluating students' schemata, interaction, purpose, thinking, and metacognition. The idea of minimum competency testing is repugnant to most. The simplistic idea that one can successfully evaluate prospective teachers by testing them for minimum knowledge and skill abilities is as ridiculous as the idea of successfully evaluating students by testing them for minimum reading subskill proficiency.

How does this new era affect the field of reading? Until recently, many in reading education did not seem to acknowledge that there would be an effect on this field or programs. However, the effect is, and will continue to be, dramatic. Reading education has often operated separately from the "real world" of the public schools and the state departments of education. Problems have been encountered in convincing the schools that reading instruction is much more than three or four ability groups doing "round-robin" reading and filling out worksheets. Little that has been taught in reading education courses appears to be implemented in the schools, possibly because little instruction seems to really influence reading education students (Gray, 1984; Robinson, Goodacre, & McKenna, 1978; Wray, 1985). This new minimum competency era will probably make it much harder to implement necessary changes in the thinking of the public schools. While preservice and inservice teachers are told that minimum skill development is not the way to proceed, the "new era" model of minimum skill testing is looming in the background. Reading education students have a hard enough time making room in their schematas for the theories espoused in university classrooms. Now with their own competencies being evaluated with the types of measures often denounced, the influence of theories espoused may be weakened further.

As a reading educator, it is important to the author to convince students to develop ways of enhancing their students' learning while allowing this learning to be stimulating and meaningful. Prospective teachers are encouraged to use different types of assignments and assessments, rather than the excessive use of worksheets and boardwork. Stimulating this type of approach is done through classroom instruction; however, it is counterproductive for these preservice teachers to know that their future evaluations will be based on the limited types of tests often cautioned against.

Many graduate students will also be asked to take and pass competency tests in reading instruction. Again, many of these tests tend to focus on discrete skills and on a tunnel-vision perspective (see Flippo & Foster, 1984 ; Flippo, Hayes, & Aaron, 1983). While university faculty try to develop more holistic viewpoints in graduate courses, students will need to grapple with minimum competency test questions with a very traditional and skill oriented focus in real life. This is a genuine problem. (See Flippo, 1986, for a full discussion of the issues and problems surrounding teacher competency testing.)

Additionally, ignoring the types of tests teacher education students must take for certification may put programs in jeopardy or under attack. Results of the tests tend to indicate that programs focusing on test objectives have more students passing the tests, and that programs not focusing on test objectives do not have pass rates as high. This is why many colleges and universities are now focusing on the test objectives and teaching courses on how to pass the teacher certification tests (see Flippo, 1986). Certainly, these tests have placed universities on the defensive (Jacobson, 1985).

Teaching the test is not advocated. However, reading education needs to stay atuned to developments in this teacher testing era. Like it or not, these tests affect students and programs. Awareness of the movement can help fight potential impact on students and their thinking. It certainly cannot hurt to be aware. The American Reading Forum acknowledges the significance of this problem by emphasizing evaluation in the theme of the 1985 annual conference. Continued discussion and interaction will be necessary to deal effectively with the impact this new era is making. This era, like other eras before it, will eventually pass. However, it behooves reading educators not to close our eyes until it does pass but to seek ways to exercise influence on emerging patterns for developing competency in effective and positive ways.

References

Bosworth, I. (1985, December). Communication from I. Bosworth, Educational Testing Service, Evanston, IL.

Flippo, R. F. (1985). Teacher certification testing across the United States and a consideration of some of the issues. Paper presented at the American Educational Research Association Annual Meeting, Chicago, IL.

Flippo, R. F. (1986). Teacher competency testing: Perspectives and issues. Journal of Teacher Education, 37.

Flippo, R. F., & Foster, C. R. (1984). Teacher competency testing and its impact on educators. Journal of Teacher Education, 35, 10-13.

Flippo, R. F., Hayes, D., & Aaron, R. L. (1983). Teacher competency testing and reading specialty preparation in Georgia. In G. McNinch (Ed.), Reading research to reading practice, Third Yearbook of the American Reading Forum. Athens, GA: American Reading Forum.

Gray, M. J. (1984). Is the influence of psycholinguistic research evident in preservice teachers' views of the reading process? Reading Horizons, 24, 253-258.

Haney, W. (1984). Testing reasoning and reasoning about testing. Review of Educational Research, 54, 597-654.

Jacobson, R. L. (1985, March). States' efforts to improve schoolteaching called inadequate. Chronicle of Higher Education, p. 20.

McCarthy, M. M. (1985, March/April). Competency tests in public employment: A legal view. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Robinson, R. D., Goodacre, E. J., & McKenna, M. C. (1978). Psycholinguistic beliefs: A cross-cultural study of teacher practice. Reading Improvement, 15, 134-137.

Sandefur, J. T. (1985). Competency assessment of teachers: The 1984 report. Unpublished manuscript.

Schnittjer, C. J., & Flippo, R. F. (1984, April). Competency based certification of school administrators: The Georgia experience. Paper presented at the 1984 American Educational Research Association Annual Meeting, New Orleans, LA.

Teacher competence. (1985). Bloomington, IN: Phi Delta Kappa's Center on Evaluation, Development and Research.

Wray, D. (1985). The impact of psycholinguistic theories on trainee-teachers' views of the teaching of reading. Unpublished manuscript.

HOW EFFECTIVE ARE QUANTITATIVE TEACHER EVALUATIONS?

BETH ANN HERRMANN
University of South Carolina

To date, instructional effectiveness is associated with the presence or absence of overt teacher behaviors (Brophy and Good, in press). Consequently, most models for evaluating and improving teachers' instructional performance emphasize quantitative recording and coding of observable behaviors. An evaluator (usually a principal) observes a teacher and systematically records on a checklist or profile sheet the presence or absence of specified teacher behaviors, their frequency and/or their duration. Qualitative data, descriptive evidence of how well the teacher performed each of the specified behaviors, are seldom a part of quantitative evaluation models.

The assumption is made that quantitative evaluations assist teachers in improving their instructional performances. Recorded information is summarized and coded and/or rated to determine the extent to which the teacher's instructional performance meets certain criteria. A conference with the teacher follows during which the evaluator highlights areas where this quantitative approach revealed the absence of one or more behaviors. Little is done to improve the quality of teacher behaviors which may be quantitatively present but in a less than optimum form.

Teacher evaluation models which focus exclusively on quantitative recording and coding of observable behavior may not help teachers improve the quality of their instruction. Results of quantitative evaluations reveal little about qualitative aspects of what teachers actually say during instruction. For example, two teachers, both of whom communicate objectives to their students, may receive high evaluations because both demonstrate the specified behavior. Qualitatively, however, one teacher's instructional statements may be more effective than the other. Such qualitative differences between teachers' instructional talk are frequently overlooked during quantitative recording and coding of observable behaviors. Teacher performances would improve if more emphasis was placed on recording and evaluating what teachers say during instruction.

The purpose of this paper is to present an alternative to the quantitative model for teacher evaluation, one which includes qualitative as well as quantitative evaluations of instructional performances. To support the need for such a model, results of quantitative and qualitative evaluations of two teachers (Teacher P and Teacher S), who each taught mathematical problem solving strategies to similar students, are presented. Both teachers received high quantitative evaluations for their instructional performances during one lesson, but only one teacher, Teacher P, had students who understood what was taught, why it was taught and how to use it, a condition attributed to the qualitative differences in the two teachers' instruction. Qualitative aspects of both teachers' instructional talk are emphasized in this paper to show what the quantitative evaluation did not reveal about each teacher's instructional performance.

Background

Data reported here are based on an analysis of data collected as part of a descriptive study of verbal interaction patterns of four clinic teachers who taught in a five-week supervised clinical practicum at a small eastern college (Herrmann, in preparation). Each teacher was taught to use techniques of explicit explanations when teaching mathematics problem solving strategies following an explicit explanation model developed by Duffy, Roehler, Meloth & Vavrus (1985). Each teacher was then observed teaching one lesson per week, during the five-week program. Each observed lesson was video and audio tape recorded and described using field notes. After the lesson, three target students from each teacher's group were taken to nearby rooms and interviewed to determine their understanding of what was taught, when to use it and how to do it. A debriefing session with the teacher was conducted immediately after each videotaped lesson to review and analyze the explanation behavior.

Audio recordings of the lessons were transcribed and rated by raters trained in the use of an explanation rating form (Roehler, Duffy, Vavrus, Putnam, Wesselman, Sivan, Book, & Meloth, in press). Ten subcategories of the instrument focus on what the teacher said to the students about the task to be learned. Each of the subcategories has an explicitness scale of 0 to 4 with a maximum obtainable score of 40.

Audio taped student interviews were also transcribed and rated by raters trained in the use of an awareness rating form (Roehler, Duffy, Vavrus, Putnam, Wesselman, Sivan, Book & Meloth, in press). The rating form focuses on students' verbal reports of what was taught, when it would be used and how to do it. Each category has a scale of 0 to 4 with a maximum obtainable score of 12.

While the overall results of the study suggest that teachers who received high explanation ratings had students who were more aware than teachers who had low explanation ratings, there were cases in which high quantitative ratings did not predict high student awareness. One of these cases is the focus of this paper.

What Quantitative Analysis Do Not Reveal

Transcripts of each teacher's lesson on how to use a strategy for solving mathematical story problems and transcripts of each teacher's study interviews were rated. Teacher P's quantitative rating for her explanation was 38 (out of a possible 40) and the average student awareness rating for three of her students was 10 (out of a possible 12). Teacher S also received a high quantitative rating for her explanation (33), but the average student awareness score for her three target students was only 4. Both teachers received high quantitative ratings, that is, specified behaviors were present, but only one teacher's students were aware of the cognitive strategy that was taught and how and when to use it.

Close examination of the lesson transcripts suggests that the reason for the difference in student awareness is that Teacher P and Teacher S talked to their students in qualitatively different ways during their lessons. They differed in the kinds of information they presented about the strategy to be learned, when to use the strategy and how to use the strategy.

Information Presented About the Strategy

Although both teachers received identical training and both talked about the same strategy during their lessons, they communicated different messages about what was to be learned. Teacher S told her students that they were learning a strategy for solving problems, but her explanation was vague and limited to certain types of problems.

Teacher S: I'm sure math problems come up a lot especially when you're trying to figure out how to cook something. At the end of today's lesson you will know how to use a strategy to help you understand cooking and food preparation problems.

In contrast to Teacher S's vague introduction of the strategy, Teacher P made specific statements about what was taught and when it would be useful.

Teacher P: Today I'm going to teach you how to figure out how to solve hard word problems that you don't understand after you have read them. The problem solving strategy that I am teaching you is something that you're going to use to solve real problems that you come across every day. You can use this strategy when you solve real problems at home or anywhere, and you can use it to solve story problems at school.

Information Presented About How to Use the Strategy

Both teachers talked about how to use the strategy, but there are noticeable differences between what Teacher S said and what Teacher P said. Teacher S emphasized a rule to be recalled, a step-by-step procedure to be memorized. Note her emphasis on procedural recall in the following lesson excerpt.

Teacher S: Try to personalize the problem. To do that, first put it in your own words. Second, try to make a picture in your head. That's what we do before we solve the problem.

In contrast to Teacher S's emphasis on procedure, Teacher P presented the strategy as an adaptive, flexible, useful strategy to solve a problem encountered in mathematics or in real life.

Teacher P: You have to develop a strategy. You have to think of a way to see whether or not you've understood the problem. You're not going to go from reading to writing problems. That's out. That's no good. That's not the best way to go about solving problems. Try to figure out what the problem is saying by closing your eyes. See if you can get a mental picture of the problem situation. See if you can visualize the problem.

Information Presented About When to Use the Strategy

Both teachers talked to their students about when the stragegy would be used, but each conveyed different information about the usefulness of the strategy. Teacher S did not emphasize immediate usefulness of the strategy and did not include a real problem in her presentation. Instead, she relied on a vague situation in which the strategy might be used.

Teacher S: I'm sure math problems come up a lot especially when you're trying to figure out how to cook something and what ingredients to put in when making something. These can often cause trouble.

In contrast, Teacher P's explanation was presented within an meaningful framework, and she emphasized immediate usefulness of her strategy for solving real life problems.

Teacher P: We have a problem here in the clinic today that we need to solve. The mother who brought our snack today didn't bring enough cookies for all the kids. Later on I want to use our strategy to solve that problem.

Student Interview Responses

Qualitative differences in the instruction of Teachers P and S were reflected in the student awareness interview responses. When asked what they had learned in the lesson, Teacher S's students typically gave vague responses that did not relate specifically to the strategy the teacher was teaching. For example, one student said, "I was learning to be better at math" and another said, "I was learning how to solve math problems." Teacher S's students also had difficulty reporting how to do the strategy, typically describing it as a set of steps or a rule to follow. They did not demonstrate any awareness that there was a reasoning process to employ. In addition, Teacher S's students were unsure of the usefulness of the strategy, reporting that the strategy would be useful in math but not when they would actually use it.

In contrast, Teacher P's students made specific statements about the strategy they had learned. One student said, "I learned something to do to help me figure out how to solve hard story problems." These students also described the strategy as a mental process. For example, one student said, "If I read a problem and I don't know how to solve it, I'll close my eyes and try to picture it in my mind. If that doesn't help I'll try to put it in my own words. Then I'll be able to figure out how to solve the problem." Teacher P's students were aware of when the strategy would be useful. One student indicated that he would use the strategy following the summer clinic to solve hard story problems in school.

Discussion

Although Teacher P and Teacher S received identical training and both demonstrated specified explanation behaviors, there were qualitative differences in what they said to their students. As a result there were noticable differences in the students' understanding of the lesson content.

Teacher P provided explicit and precise statements about the strategy, why it was important, when it would be useful and how to use it. In addition, she modeled the mental reasoning employed when using the strategy. When students misunderstood how to use the strategy, Teacher P responded directly to their restructuring of the information. Teacher P's student interview responses indicated a high level of student understanding of what was taught, why it was taught, when it would be useful and how to use it.

On the other hand, Teacher S was less explicit and precise about the strategy, why it was important, when it would be useful and how to do it. Instead, she presented and modeled the strategy as a rigid set of steps to employ when solving problems. For the most part, Teacher S interrogated students to determine their understanding (or misunderstanding) of the strategy and rarely elaborated on students' responses or guided their restructuring of information presented. Teacher S's student interview responses indicated a lack of student understanding of what was taught, why it was important and how to use it.

Quantitative ratings of Teacher P's and Teacher S's instructional performances indicated that both demonstrated specified behaviors associated with effective explanations. Qualitative evaluations, however, revealed important additional information about the teachers' conceptual interpretations of the mental reasoning associated with the strategy. These qualitatively different conceptual understandings influenced each teacher's instructional talk which, in turn, influenced the students' understanding. These results have important implications for improving teacher evaluation procedures currently employed in most public school districts.

An Alternative to Quantitative Teacher Evaluation Models

Results of qualitative evaluations of teachers' instructional talk suggest that even when expected behaviors are present, noticeable differences in student understanding result if there are qualitative differences in teachers' instructional talk. Therefore, teacher evaluation models must focus not only on overt teacher behaviors but also on qualitative dimensions of what teachers say during instruction and the relationship between teachers' instructional talk and student outcomes.

Implicit in any model of teacher evaluation is a shift in the role of the evaluator from observer to instructional leader. In addition to receiving training to record and code information about expected teacher behaviors, instructional leaders must be trained to identify effective and less effective teacher statements. Observed lessons should be tape recorded and reviewed during a series of debriefing sessions designed to help less effective teachers plan better lessons and make qualitatively better statements as they teach.

Conclusion

Quantitative teacher evaluation models may adequately identify effective and less effective teachers, but they reveal little about qualitative dimensions of what teachers say during instruction. Students' understanding of what is taught, why it is important and how to perform certain tasks is directly related to what teachers say during instruction. It is crucial, therefore, that teacher evaluation models focus on qualitative differences between teachers' instructional talk as well as quantitative recording and coding of specified behaviors. Furthermore, the role of the evaluator must shift from observer to instructional leader. Only then will the teacher evaluation assist teachers in improving their instructional performances.

REFERENCES

Brophy, J., & Good, T. (in press). Teacher behavior and student learning. In M. C. Wittrack (Ed.) _Third Handbook of Research on Teaching_. Chicago: Rand McNally.

Herrmann, B. A. (in preparation). _Strategic Problem Solving of Mathematical Story Problems: A Descriptive Study of the Effects and Characteristics of Direct Teacher Explanation_. A paper presented at the annual conference of the American Educational Research Association, San Francisco, April, 1986.

Duffy, G., Roehler, L., Meloth, M., & Vavrus, L. (1985). _Conceptualizing instructional explanation_. A paper presented at the annual conference of the American Educational Research Association, Chicago.

Roehler, L., Duffy, G., Vavrus, L., Putnam, J., Wessleman, R., Sivan, E., Book, C., & Meloth, M. (in press). _A descriptive study of teacher explanation: Final report of the 1983-84 study_. East Lansing: Institute for Research on Teaching, Michigan State University.

AN INVESTIGATION OF THE RELATIONSHIPS BETWEEN STUDENT ACHIEVEMENT AND TEACHERS' PERCEPTIONS OF INSTRUCTIONAL HINDRANCES IN THE ELEMENTARY SCHOOL

DONALD D. NEVILLE
KATHERINE RICHMOND
EVELYN F. SEARLS
University of South Florida

The time-on-task research of the last decade has clearly revealed that time allocated for instruction and time actually engaged in learning activities with high success rates are major factors in student achievement (Brophy, 1974, 1979; Fisher & Berliner, 1985; Karweit, 1981; Stallings, 1976). One result of the focus on the importance of school time has been the investigation of factors that either directly or indirectly erode the amount of time available for instruction. A direct negative effect may be seen in the large percentages of classroom time found to be devoted to administrative procedures and contending with interruptions (Conant, 1973; Denham, 1980; Karweit, 1981). Stallings, Needels, and Stayrook (1979) found that classrooms having more intrusions from the outside (e.g., announcements on the intercom, requests for students to leave the room, tardy students coming into the room) had students who made less gain in basic reading skills. An indirect effect was confirmed by Stallings and Mohlman's (1981) finding that in schools where teachers perceived more administrative hindrance (e.g., excessive record keeping, extra routine duties, and too many committee requirements), teacher morale was lower and there was more student misbehavior.

In 1984 Searls, Neville, and Richmond used a questionnaire to survey a large random sample of elementary teachers as to the degree to which the teachers perceived that their capabilities to instruct was limited by a variety of factors present in their classrooms. Although some of the factors impinged directly on teaching effectiveness and efficiency, and others indirectly, Searls et al. (1984) did not provide a comparison of the two types of effects, per se. They reported that classroom interruptions, student related problems, and class size were ranked high as hindering factors. However, when they analyzed their data by school, there were significant differences among schools, indicating that teacher-perceived hindrances were greater in some schools than in others.

The current study is an extension of previous study (Searls et al., 1984). One school factor that might affect teachers' perceptions of hindrances is student achievement levels; for example, classroom interruptions might be perceived as a greater limitation by teachers of low achieving students than by teachers of high achieving students. Therefore, the purpose of this study was to explore the possible relationships between teachers' perceptions of both direct and indirect instructional hindrances and student achievement in reading and mathematics. Because of the high relationship frequently demonstrated between achievement and socioeconomic status (SES), it seemed advisable to minimize the effect of SES by means of separate analyses. Thus, our research questions were:

1. Will there be a significant difference between the perception of both direct and indirect instructional hindrances by teachers in low SES low achieving schools as compared to such perceptions by teachers in low SES high achieving schools?

2. Will there be a significant difference between the perception of both direct and indirect instructional hindrances by teachers in a high SES low achieving schools as compared to such perceptions by teachers in high SES high achieving schools?

Method

Sample Selection

The sample selection occurred in two stages. This was necessary because the achievement and SES data for any school year were not available from the county until after that school year had ended (i.e., during the summer).

Initial stage. During the summer of 1984 approximately 45% of the highest and the lowest 45% of the elementary schools (grades K-5) in a large metropolitan county were identified according to ranking in combined reading and mathematics achievement and ranking in SES. Achievement data was based on the school means on the county standardized achievement test administered in the spring of 1984. SES was based on the percentage of students within schools who received free and reduced-price lunches during the 1983-84 year. This sampling provided a list of schools to which the questionnaire was sent in the spring of 1985.

Final stage. After achievement and SES data for the 1984-85 year became available in the summer of 1985, the researchers redefined the sample, selecting 35% of the highest and the lowest 35% of the elementary schools according to ranking in combined reading and mathematics achievement and ranking in SES. This sampling of elementary schools provided the data that was analyzed in order to answer our research questions.

Procedures

Construction of the questionnaire. The questionnaire devised by Searls et al. (1984) was used with some revisions. Some items were reworded, a few items requesting demographic information were added, and the rating scale was changed from five to seven points. (A copy of the revised questionnaire is included at the end of this article.)

Data collection. In April 1985 a packet was sent through the school mail to the principal of each elementary school identified in the initial stage of the sample selection. The packet contained a cover letter to the principal describing the study and the nature of the information requested. Included in the packet were sufficient teacher cover letters, questionnaires, and return envelopes for all classroom teachers in each school. Each teacher was asked to respond to the questionnaire anonymously, to seal it in an envelope provided, and to give it to the school secretary. The secretary then forwarded these sealed envelopes in a packet through the school mail to the county research office. After all schools had responded, the investigators collected the packets from the county office.

The percentage of teachers responding to the questionnaire ranged from 21% in one school to 100% in three schools; overall 51% (*n* = 728) of the teachers completed the questionnaire. Thus, a second mailing would have been desirable. However, permission for this was denied by the county because of concern about excessive demands on teacher time at the end of the school year.

Data analysis. The final stage of the sample selection resulted in the formation of four groups of schools: Group I, low SES and low achievement (L,L); Group II, low SES and high achievement (L,H); Group III, high SES and low achievement, H,L); and Group IV, high SES and high achievement (H,H). The correlation of achievement with SES was .73 for this sample; this resulted in smaller *n*'s for Groups II and III. (See Table 1.)

In order to minimize the effect of SES, separate analyses by SES were conducted; that is, comparison was made between Group I to Group II (both low SES) and Group III to Group IV (both high SES). In addition to identifying individual questionnaire items that discriminated between groups, group mean scores were compared on: (a) all 12 questionnaire items, (b) the 7 questionnaire frequency items, (c) items that appeared to reflect direct hindrance of instruction (questionnaire items H1 + H2 + H3 + H5 + H9 + H10), and (d) items that appeared to reflect indirect hindrance of instruction (questionnaire items H4 + H6 + H7 + H8 + H11 + H12).

Table 1

Questionnaire Subtotals by Group

Group [a]	*M*	*SD*	*n*	*t*	*df*	*p*
			Hindrances (HT3) [b]			
I	46.85	12.08	162			
II	38.85	10.09	27	3.704	187	< .01
III	45.57	10.66	47			
IV	42.05	12.06	158	1.913	203	NSD
			Frequencies (FT1) [c]			
I	31.30	6.55	174			
II	31.26	7.49	27	.020	199	NSD
III	30.89	4.96	47			
IV	31.10	6.22	174	-.300	219	NSD
			Direct Effects (HT1) [d]			
I	25.32	6.56	178			
II	21.78	6.06	27	2.146	203	< .05
III	25.04	5.97	48			
IV	22.95	6.64	168	2.885	214	< .01
			Indirect Effects (HT2) [e]			
I	21.36	6.49	172			
II	17.07	5.03	27	2.134	197	< .05
III	20.33	5.41	48			
IV	19.23	6.51	175	1.581	221	NSD

[a] Group I = low SES, low achievement; Group II = low SES, high achievement; Group III = high SES, low achievement; Group IV = high SES, high achievement. [b] HT3 = H1 - H12 (items under Hindrance Column on Questionnaire). [c] FT1 = F1 - F7 (items under Frequency Column on Questionnaire). [d] HT1 = H1 + H2 + H3 + H5 + H9 + H10. [e] HT2 = H4 + H6 + H7 + H8 + H11 + H12.

Results

Characteristics of the Sample

Respondents were asked to provide data concerning age, years of teaching experience, sex, type of assignment, and grade(s) taught. Information was also requested concerning whether the students were taught all day by the same teacher or cross-grouped for reading and/or mathematics; however, the grouping data were not analyzed since less than 60% of the initial sample provided interpretable responses.

Numbers of respondents specifying the information requested varied according to the nature of the information. In the initial sample ($\underline{n}$ = 728), percentages of respondents specifying age, years of teaching experience, sex, type of assignment, and grade(s) taught ranged from 84% to 97%. In the final sample ($\underline{n}$ = 470), percentages specifying ranged from 82% to 97%. The pattern of responses among the four comparison groups were substantially similar.

The final sample appeared to be reflective of the initial sample in mean age, mean length of teaching experience, and proportional representation by sex, type of assignment, and grade(s) taught. Moreover, the four comparison groups were both similar to each other and representative of the final sample on these factors.

Based upon the total of respondents specifying the requested information, the mean age of this group was 38.8 years; mean length of teaching experience was 12.4 years. Of those respondents specifying sex, 9% were male and 91% were female. The sample consisted of regular classroom teachers (72%); music, art, or physical education teachers (8%); and "other" teachers (20%) who were generally engaged in some form of special education, remediation (reading, Chapter I), or special services (guidance, speech, media). Of those teachers specifying grade(s) taught, 14% were kindergarten teachers; 39%, primary teachers (grades 1, 2, 3, or a combination thereof); 27%, intermediate (grades 4, 5, or a combination thereof); 2%, primary/intermediate combination; and 19% indicated their assignments went across grades K-5.

School SES Groups and Individual Questionnaire Items

Groups I and II. As stated previously, these two groups both represented low SES schools, but Group I (L,L) was comprised of low achieving schools and Group II (L,H) of high achieving. A chi-square analysis was applied comparing the responses of the two groups on each questionnaire item. Since single item scores were used as the dependent variable, a $p<.10$ value was selected as the indi-cator of significance.

The items that significantly discriminated were: F1, whole class interruptions (e.g., loud speaker); F7, communication problems, H1, same as F1 except degree of hindrance rather than frequency; H8, physical aspects (e.g., space and equipment); and H10, individual student problems. In all instances Group I responded with higher scores than Group II.

Groups III and IV. In this comparison both groups represented high SES with Group III (H,L) being low achieving and Group IV (H,H), high achieving. Again, a chi-square analysis was applied to each questionnaire item using a $p<.10$ as the criterion for significance.

The items which discriminated were: F3, whole class interruptions, (e.g., assembly or special program); H8, physical aspects (e.g., space); H9, large classes; and H10, individual student problems. In all instances Group III responded with higher scores than Group IV.

Synthesis. If one considers the relationship of the results of the separate group comparisons, there are some interesting observations. In all items identified as discriminating, the low achieving Groups (I and III) rated the items higher than did the high achieving comparison groups. Also, there were two items, H8 and H10, common to both comparisons, and F1 and F3 were both whole class interruptions even though they occurred for different reasons.

School SES Groups and Questionnaire Subtotals

Groups I (L,L) and II (L,H). All comparisons were made using a t test for mean differences Since each score was comprised of six or seven items, a $p<.05$ level was the criterion for significance.

The HT3 subtotal (all degree of hindrance items) showed a significant difference, with Group I being higher. The FT1 (frequency) subtotal difference was not significant. The HT1 subtotal, consisting of the six items representing direct hindrances, showed a significantly higher mean score for Group I. The HT2 subtotal, consisting of the six indirect hindrance items, showed a significant difference, with Group I being higher. (See Table 1.)

Group III (H,L) and Group IV (H,H). Again, all comparisons used t tests and a $p<.05$ criterion. No significant differences were found for three of the four comparisons. Only one subtotal, HT1 (direct hindrances), was identified, with the Group III mean score being significantly higher. (See Table 1.)

Synthesis. For these eight comparisons, four were identified as discriminating significantly between the groups (either I and II or III and IV). In all instances, the lower hindrance scores were made by the high achieving groups In addition, of the four significant differences, three were related to Groups I and II (both low SES). Lastly, there were no indications that any group saw the hindrances occurring more frequently than any other group.

Conclusions

Limitations

1. As with any questionnaire, the sample consisted of volunteer respondents; therefore, the response rate of 51% was somewhat lower than one would hope.

2. The constitution of these school groups, resulting in very large differences in the number of responses per group being analyzed, led the researchers to an analysis which used a sequence of chi-squares and t tests. This may have resulted in some chance significant differences.

3. Of course, the evidence from these data were relationships between teacher perceptions of hindrances and achievement level. There is no foundation for the argument that the relationship is causal.

Achievement and Hindrances

It appears that teachers in low achieving schools clearly saw hindrances as having a more serious effect on instruction than did teachers in high achieving schools. This occurred even though their perception of the frequency of occurrence of these hindrances did not differ.

It also appears that teachers in low achieving low SES schools perceived more of these hindrances as having a serious effect on instruction that did any other group. Group I found five individual items and the HT1, HT2, and HT3 subtotals as having a more negative effect. In contrast, Group III found four individual items and only the HT1 subtotal as having a more negative effect. Thus, high SES seemed, in the teachers' view, to reduce the negative effect of hindrances on instruction.

The results of this research appear consistent with earlier findings as reported in the literature. Four studies reported that, when large percentages of classroom time were devoted to administrative procedures and/or contending with interruptions (i.e., direct hindrances), achievement was negatively affected (Conant, 1973; Denham, 1980; Karweit, 1981; Stallings et al., 1979). In addition, negative effects for indirect hindrances were found by Stallings and Mohlman (1981).

No information explained why the relationship between achievement and hindrances occurred. Perhaps future research could explore causes. Interviews might be used to further investigate why teachers responded the way they did. Another approach might be to select some other variables, for example, teacher anxiety or morale (Stallings & Mohlman, 1981), and investigate the relationship to teaching in schools representing different SES and/or achievement levels.

QUESTIONNAIRE: FACTORS HINDERING CLASSROOM INSTRUCTION

School ______________________ Regular classroom teacher ______________
Grade(s), 1984-85

Other than regular classroom teacher: ______________ ______________
Title Grade(s), 1984-85

Frequency	Degree of Hindrance
1=Never happens	1=Never hinders instruction
2=Happens less than once per semester	2=Very seldom hinders instruction
3=Happens less than once per month	3=Seldom hinders instruction
4=Happens less than once per week	4=Sometimes hinders instruction
5=Happens once per week	5=Often hinders instruction
6=Happens two-four times per week	6=Very often hinders instruction
7=Happens five or more times per week	7=Always hinders instruction

Items 1-7: In the first column (Frequency) please indicate how often each of the following occurs. In the second column (Degree of Hindrance) please indicate the extent to which each of the following hinders your instruction of students. Circle the appropriate number in each column. Please react to each item in an overall manner; examples are given only for clarification.

	Frequency	Degree of Hindrance
1. All interruptions which disturb the whole class (e.g., loudspeaker, messenger, request for information/equipment, etc.)	1 2 3 4 5 6 7	1 2 3 4 5 6 7
2. All interruptions of one or more students leaving the room or returning (e.g., called to office, involved in SLD, Chapter 1, Chorus, newspaper, etc.)	1 2 3 4 5 6 7	1 2 3 4 5 6 7
3. All interruptions when whole class is scheduled for special programs (e.g., assembly, movie, middle school orientation, etc.)	1 2 3 4 5 6 7	1 2 3 4 5 6 7
4. All required testing and related record keeping (e.g., state assessment, county-wide testing, pre-post skill testing, etc.)	1 2 3 4 5 6 7	1 2 3 4 5 6 7
5. All system- and/or other school-required record keeping (e.g., book count, money collection, IEP's, etc.)	1 2 3 4 5 6 7	1 2 3 4 5 6 7
6. All miscellaneous duties (e.g., typing and/or duplicating materials, cafeteria and/or bus duty, meetings unrelated to instruction, etc.)	1 2 3 4 5 6 7	1 2 3 4 5 6 7
7. All communication problems (e.g., those among teachers, between teachers and administrators, between classroom & resource teachers, etc.)	1 2 3 4 5 6 7	1 2 3 4 5 6 7

Degree of Hindrance

1=Never hinders instruction
2=Very seldom hinders instruction
3=Seldom hinders instruction
4=Sometimes hinders instruction
5=Often hinders instruction
6=Very often hinders instruciton
7=Always hinders instruction

Items 8-12: For the following items indicate the Degree of Hindrance only.	Degree of Hindrance
8. All aspects of the physical environment (e.g., noise; inadequate or inappropriate space, equipment, materials, or facilities, etc.)	1 2 3 4 5 6 7
9. All problems occurring as a result of large class size (e.g., discipline, wide disparity in achievement levels, etc.)	1 2 3 4 5 6 7
10. All circumstances related to individual student problems (e.g., discipline, lack of parental support, poor physical and/or mental health, high or low achievement and/or ability level very different from class average, etc.)	1 2 3 4 5 6 7
11. All curriculum-related problems (e.g., lack of time to cover mandated content, lack of flexibility, lack of teacher input, curriculum determined by tests, special programs such as drug or guidance programs, etc.)	1 2 3 4 5 6 7
12. All administrative support problems (e.g., not backing teacher actions with students and parents, not providing flexibility in budget allocation, not involving teachers in school planning, etc.)	1 2 3 4 5 6 7

At your grade level, do you teach the same group of students for all subjects?
Yes_____ No _____ Not Applicable _____

At your grade level, are students cross-grouped for reading instruction?
Yes_____ No _____ Not Applicable _____

At your grade level, are students cross-grouped for mathematics instruction?
Yes_____ No _____ Not Applicable_____

Your age ______ Sex ___ Years of Teaching Experience _____

Thanks for your help. Please seal this questionnaire in the envelope in which it was received and give to your school secretary for forwarding to the county research office by April 23rd.

References

Brophy, J. E. (1974). Achievement correlates. In H. J. Walberg (Ed.), Evaluating educational performance. Berkeley, CA: McCutchan.

Brophy, J. E. (1979). Teacher behavior and its effects. Journal of Educational Psychology, 71, 733-750.

Conant, E. H. (1973). Teacher and paraprofessional work productivity: A public school cost effectiveness study.

Denham, C. H. (1980). Lessons for Policymakers. In C. Denham & A. Lieberman (Eds.), Time to learn. Washington, DC: Department of Education (National Institute of Education).

Fisher, C. W., & Berliner, D. C. (Eds.) (1985). Perspectives on instructional time. New York: Longman.

Karweit, N. (1981). Time in school. Research in sociology of education and socialization, 2, 77-110.

Searls, E. F., Neville, D. D., & Richmond, K. (1984). Teachers' perceptions of factors that hinder reading instruction. In G. H. McNinch (Ed.), Reading teacher education (4th Yearbook of the American Reading Forum). Athens, GA: University of Georgia.

Stallings, J. (1976). How instructional processes relate to child outcomes in a national study of Follow Through. Journal of Teacher Education, 27, 43-47.

Stallings, J., & Mohlman, G. (1981). School policy, leadership style, teacher change, and student behavior in eight schools. Final report to the National Institute of Education. Menlo Park, CA: SRI International. (ERIC Document Reproduction Service No. ED 209 759)

Stallings, J., Needels, M., & Stayrook, N. (1979). How to change the process of teaching basic reading skills in secondary schools. Final report to the National Institute of Education. Menlo Park, CA: SRI International. (ERIC Document Reproduction Service No. ED 210 671)

REACTION PAPER

Commentary on AN INVESTIGATION OF THE RELATIONSHIPS BETWEEN STUDENT ACHIEVEMENT AND TEACHERS' PERCEPTIONS OF INSTRUCTIONAL HINDRANCES IN THE ELEMENTARY SCHOOL by Donald D. Neville, Katherine Richmond and Evelyn F. Searls

Reactor: MARINO C. ALVAREZ
Tennessee State University

Neville, Richmond, and Searls identify a pertinent issue that confronts classroom teachers everyday. As they indicate, varying hindrances to instruction are a common occurrence that affect teachers and students in various ways. Their study shows that teachers of low socioeconomic (SES) students and those in low achieving schools viewed these hindrances as major obstacles to their instruction.

Schools were selected based upon rankings of combined reading and mathematic scores on a county standardized achievement test. Teachers were asked to respond to a mail questionnaire by rating the degree and frequency of an hindrance on a Likert type scale. The questionnaire contained twelve items. The first seven statements required participants to rate the frequency and degree of hindrance, and to "react in an overall manner; examples are given only for clarification." The five statements that remained, dealt with degree of hindrance only. Neville et al., mention the commonality of item 8, an indirect hindrance, (aspects of the physical environment) and item 10, a direct hindrance (individual student problems), and items 1 (whole class interruptions) and 3 (whole class interruptions scheduled for special programs) in degree of frequency.

Teachers in low SES and low achieving schools differed from teachers in low SES and high achieving schools in perceived hindrances caused by aspects of the physical classroom environment and degree of individual student problems. Teachers in high SES and low achieving schools differed from teachers in high SES and high achieving schools by responding with higher rating to items dealing with special programs involving entire class in the degree of frequency, the degree of hindrance caused by aspects of the physical classroom, large classes, and individual student problems. Low achieving schools, rather than socio-economic status, were the common denominator.

It may be that because the statements were matched with specified examples that the data interpretation was confounded. The respondents may have been influenced by the provided examples.

Also, there are nested variables within statements as well as overlapping examples (i.e., "discipline" in 9 and 10 both classified as direct instructional hindrances). It is not clear whether teachers responding to the statements scored a statement high or low on the basis of what they read as a category, an example, or if they perceived another situation under the stated category that was not listed among the examples as the primary hindrance.

Data interpretation is further complicated by a low questionnaire response rate of 51% and the absence of information about why teachers responded as they did.

In this study, it would have been interesting to determine what influenced teacher choice of items. The reasons why teachers of low achieving schools perceived hindrances seriously affecting instruction more so than teachers in high achieving schools might have been obtained by an interview of a random sample of the participating teachers. An interview would have clarified why teachers believed certain practices hindered their instruction. The nested variables contained within each item could be separated precisely through interviews with participants. It may be that items classified as indirect may, in fact, be perceived as direct causes by some teachers. A reason why high achieving schools did well on the county test might have been a function of how these schools prepared their students as opposed to low achieving schools.

Why high achieving schools were not bothered to the extent of low achieving schools may be idiosyncratic to a particular school operation. It may be interesting to compare the responses from the three schools that reported 100% returns with those that numbered low returns. A follow-up interview may reveal that the school environment may be the key factor overriding socio-economic influence.

Teacher perceptions could further be substantiated or refuted by conducting ethnographic research to determine the degree to which perceived hindrances are in essence direct causes of instruction and learning problems or are unfounded perceptions. It may be that the perceptions were idiosyncratic to schools rather than determined by low achieving or high achieving variables.

The study is noteworthy in that these researchers have raised questions that need to be investigated further. Teacher perceptions affect student and school performance. These perceptions need to be investigated and resolved. Teachers and administrators need to be aware of the degree to which hindrances are affecting school learning so that, once documented, they can be either curtailed or avoided.

EVALUATION OF A READING PROGRAM: TEACHER JUDGMENT AND PUPIL PLACEMENT

BETH S. WISE
JUDY C. SHAVER
McNeese State University

In order to evaluate a reading program, it is important to go beyond numbers being served and reading test scores, to examine the population being served in relation to the needs of that population. This study examined the sex and race of students of the population referred as compared to those served in a special reading program. It also examined the effect of teacher judgment in the referral process.

Background

Past research indicates that, in general, boys have more difficulty in learning how to read than girls. This is especially true in the lower grades. Recent studies have examined cultural influences and sex role expectations as possible causal factors. They point to cultural factors such as teacher expectation that girls will perform better in reading, to the social structure of our elementary schools, to the importance of preschool education programs, and to parents' views of sex-role stereotypes and the roles they play in expectations (Downing, 1975; Johnson, 1976; Downing & Thomson, 1977; Gross, 1978; Shapiro, 1980). While boys are designated for special reading classes by margins of from 1:1 to 10:1, research indicates that test score differences between boys and girls may not be as great as was formerly thought (Naiden, 1976).

One possible explanation for boys being disproportionately represented in special reading classes could be the presence of sex bias in the referral process. If this is so, girls may not be receiving special help in reading in proportion to their needs (Thorndike, 1976; Johnson & Greenbaum, 1980).

Purpose of the Study

This study had the following purposes: to examine the role of teacher judgment in the placement of pupils in a special reading program, to examine the ratio of black and white boys and girls referred to those being served, and to determine whether either sex or race was disproportionately represented.

Procedures

Data were gathered on students in grades 2, 3, and 4 in 21 public and 4 parochial elementary schools served by Chapter One in Calcasieu Parish in southwest Louisiana. Data were collected on those students who were considered to be eligible for special reading classes in the school years 1980-81 and 1981-82. The sample included a total of 1,380 boys and 978 girls. The descriptive data collected on each subject included grade level, sex, ethnic origin, SRA reading score, Teacher Judgment score, and assignment to Chapter One reading classes in Calcasieu Parish.

Pupils were assigned to this federal program based on total points accumulated using the following criteria: Teacher Judgment Score (0-18), SRA Test Score (0-18), Reading Proficiency Test (0-6), Grade Retention (6), and Past Participation in a Chapter One Program (6). At the time of the research study, particular attention was focused on the number of points assigned under the unspecified "Teacher Judgment" section of the form, as compared to SRA scores.

In order to determine the effect of teacher judgment in the referral process, data were gathered on students who fit the following criteria: SRA scores did not indicate a need for referral. (Scores were above the 41st percentile and students were assigned points of 14 or less on the eligibility form.) Teacher Judgment scores regarding "Degree of Need" did indicate a need for referral. (Students received 11 points or more on this section of the eligibility form.)

Results

The findings for Chapter One Reading by grades for 1981 are reported in Table 1. Sixty-five percent of students referred for remedial reading were boys. Of those referred, 75% were served. Compared to this, 44% of students in 1981 were girls, and 71% of the girls referred were served. There were almost equal numbers of Blacks (610) and Whites (608) referred for remedial reading. Seventy-six percent of the black students referred were served, while only 71% of the white students referred were served.

A significant difference ($p<.05$) appeared at the second grade level between referred boys served (80%) and referred girls served (66%). Eighty percent of both black and white boys referred were served. However, 70% of black girls were served, while only 61% of white girls referred were served. At third grade, black girls were served at a higher rate (87%) than others. (Black boys, 79%; white

Table 1

Chapter One Reading by Grades, 1981

	Grade 2			Grade 3			Grade 4		
	R	S	%	R	S	%	R	S	%
Black boys	150	120	80	72	57	79	134	97	72
White boys	133	106	80*	91	70	77	108	68	63
Black girls	102	71	70	55	48	87	97	70	72
White girls	98	60	61*	66	50	76	112	77	69
Total boys	283	226	80*	163	127	78	242	165	68
Total girls	200	131	66*	121	98	81	209	147	70
Total Blacks	252	191	76	127	105	83	231	167	72
Total Whites	231	166	72	157	120	76	220	145	66

*p<.05   R = Referred   S = Served

boys, 77%; white girls, 76%) However, this difference did not reach significance.

Numbers of students referred to and served by Chapter One Reading in 1982 are reported in Table 2. At fourth grade, a significant difference (p<.05) appeared between referred white boys served (64%), as compared to black boys (80%), and black girls (83%). There was also a significant difference (p<.05) at fourth grade level between total numbers of referred Blacks served (81%) and referred Whites served (69%).

In 1982, 61% of students referred for remedial reading were boys, and of these referred, 77% were served. Compared to this, 39% of students referred were girls, and of that number, 78% were served. A significant difference (p<.05) was found between the races: 82% of referred Blacks were served, whereas 73% of referred Whites were served. A significant difference (p<.05) also appeared between black and white boys: 82% of referred black boys were served, while only 70% of referred white boys were served. These findings are shown in Table 3.

Table 2

Chapter One Reading by Grades, 1982

	Grade 2			Grade 3			Grade 4		
	R	S	%	R	S	%	R	S	%
Black boys	124	96	72	135	121	90	132	105	80*
White boys	116	78	67	94	76	81	91	58	64*
Black girls	81	58	72	65	58	89	83	69	83*
White girls	89	60	67	65	57	88	65	49	75
Total boys	240	174	73	229	197	86	223	163	73
Total girls	170	118	69	130	115	89	148	118	80
Total Blacks	205	154	75	200	179	90	215	174	81*
Total Whites	205	138	67	159	133	84	156	107	69*

*p<.05 R = Referred S = Served

Table 3

Chapter One Reading Totals

	1981			1982		
	R	S	%	R	S	%
Black boys	356	274	77	391	322	82*
White boys	332	244	74	301	212	70*
Black girls	254	189	74	229	185	81
White girls	276	187	68	219	166	76
Total boys	688	518	75	692	534	77
Total girls	530	376	71	448	351	78
Total Blacks	610	463	76	620	507	82*
Total White	608	431	71	520	378	73*

*p<.05 R = Referred S = Served

In 1981, significant differences (p<.05) were found at second grade between total number of boys served (88%) as compared to total number of girls served (74%) as a result of teacher recommendation. Also in 1981, a significant difference (p<.05) appeared at fourth grade level between Blacks served (79%) and Whites served (64%) on teacher recommendation. These findings are shown in Table 4.

Table 4

Chapter One Reading by Grades, 1981
Students Referred by Teacher Judgment

	Grade 2			Grade 3			Grade 4		
	R	S	%	R	S	%	R	S	%
Black boys	48	41	85	24	18	75	40	30	75
White boys	43	39	91	44	38	86	50	31	62
Black girls	36	27	75	17	16	94	33	28	85
White girls	33	24	73	32	28	88	37	25	68
Total boys	91	80	88*	68	56	82	90	61	68
Total girls	69	51	74*	49	44	90	70	53	76
Total Blacks	84	68	81	41	34	83	73	58	79*
Total Whites	76	63	83	76	66	87	87	56	64*

*p<.05 R = Referred S = Served

Table 5 shows the findings for teacher evaluation data for 1982. No significant differences within grades were found in this data.

Discussion

In reflecting on the numbers of students both referred and served in special reading classes in Calcasieu Parish in 1981 and 1982, girls were definitely underrepresented. Blacks were served at a higher percentage when comparing students by sex. Black boys were referred in greater numbers, and were served at a higher percentage rate across the board than were white boys, black girls, or white girls. White girls were referred for remedial reading in equal or greater numbers than white boys and black girls.

Table 5

Chapter One Reading by Grades, 1982
Students Referred by Teacher Judgment

	Grade 2			Grade 3			Grade 4		
	R	S	%	R	S	%	R	S	%
Black boys	13	11	85	27	22	81	10	6	60
White boys	23	19	83	35	28	80	14	11	79
Black girls	11	10	91	12	10	83	5	4	80
White girls	15	11	73	28	25	89	8	6	75
Total boys	36	30	83	62	50	81	24	17	71
Total girls	26	21	81	40	35	88	13	10	77
Total Blacks	24	21	88	39	32	82	15	10	67
Total Whites	38	30	79	63	53	84	22	17	77

*p<.05 R = Referred S = Served

However, a smaller percentage of white girls were served in special reading classes across the board.

Looking at the effect of teacher evaluation on referrals, more white boys were referred than black boys, black girls or white girls, at all levels except in second grade in 1981, when more black boys were referred. Percentages of those referred who were served did not favor either sex or race. Percentages of those served who were referred because of teacher evaluation were high.

In 1982, the numbers of students served in the remedial reading program as a result of teacher recommendation dropped drastically. In 1981, 36% of boys served were there as a result of teacher recommendation; in 1982, only 18% fit the category. Data on girls were similar: in 1981, 35% served were there as a result of teacher judgment; in 1982, the number was 18%.

Racially, findings were similar. In 1981, 32% of Blacks and 39% of Whites were there as a result of teacher recommendation. In 1982, these percentages dropped to 13% of Blacks and 24% of Whites being referred as a result of teacher recommendation. Since referral forms were the same for these two years, no reason for this variance is apparent.

Summary, Implications

In summary, this study examined the role of teacher judgment in the pupil placement process, and the ratio of black and white boys and girls referred to those served in an attempt to determine whether either sex or race was disproportionately represented in special reading programs in Calcasieu Parish.

The significant findings ($p<.05$) of the study regarding teacher judgment included: (a) More referred boys than referred girls were served at second grade in 1981. (b) More referred black students than referred white students were served at fourth grade in 1981.

Other significant findings were: (a) More referred boys than referred girls were served in 1981 at grade two. The group receiving the least amount of service were the white girls. (b) In 1982, at fourth grade level, a greater number of referred black boys as compared to referred white boys were served. (c) In 1982, at fourth grade level, a greater number of referred black students were served than referred white students served. (d) In 1982, more referred black boys were served than referred white boys. (e) The total number of referred black students served was significantly greater than the total number of referred white students served in 1982.

Since these findings show that both boys and Blacks are disproportionately represented in special reading programs in Calcasieu Parish, teachers should be made aware of possible sex and racial biases in the referral process.

One outgrowth of the study has been revision in the Calcasieu Parish Chapter One student eligibility form. The Teacher Judgment section of the form was changed from a nonspecific "Degree of Need" section to one that includes assigned points for specific areas of reading deficiency.

A follow-up study should be conducted in order to determine whether changes in the eligibility form have had an impact on student referrals since 1982, and whether possible biases regarding sex and race have been affected.

References

Downing, J. (1975). An application of the comparative method to a practical educational problem: Literacy learning. School Review, 83, 449-459.

Downing, J., & Thomson, D. (1977). Sex role stereotypes in learning to read. Research in the Teaching of English, 11, 149-155.

Gross, A. D. (1978). The relationship between sex differences and reading ability in an Israeli kibbutz system. In D. Feitelson (Ed.), Cross-cultural perspectives on reading and research (pp. 72-87). Newark, DE: International Reading Association.

Johnson, C. S., & Greenbaum, G. R. (1980). Are boys disabled readers due to sex-role stereotyping? Educational Leadership, 27, 492-496.

Johnson, D. D. (1976). Cross-cultural perspectives on sex differences in reading. The Reading Teacher, 29, 747-752.

Naiden, N. (1976). Ratio of boys to girls among disabled readers. Reading Teachers, 29, 460-462.

Orlow, M. (1976). Literacy training in West Germany and the United States. Reading Teacher, 29, 439-442.

Shapiro, J. E. (1980). Primary children's attitudes toward reading in male and female teachers' classrooms: An exploratory study. Journal of Reading Behavior, 12, 256-257.

Thorndike, R. L. (1976). Reading comprehension in fifteen countries. In J. Merritt (Ed.), New horizons in reading (pp. 500-507). Newark, DE: International Reading Association.

STUDENTS' PERCEPTIONS OF THEIR STUDY HABITS

BRENDA S. TOWNSEND

The Pennsylvania State University

The successful college student learns more than just the content of their courses; they also learn how to enhance their own knowledge. They learn how to learn by developing effective independent learning strategies (Brown, Campione & Day, 1981). Historically, developmental courses have been designed to help students learn effective study strategies. Developmental educators employ several techniques to meet this challenge. However, the effectiveness of these strategies is questionable unless the students choose to change traditional habits and utilize these strategies.

Students tend not to use a wide variety of effective learning strategies. That conclusion has been drawn by several researchers. Tierney, LaZansky, and Schallert (1982) found that students rarely varied their study strategies. Their strategies included rote memorization of parts of the text (especially those that provided the answers to teacher prepared questions) and the completion of any assigned activities. Simpson (1983) also found that students would continue to use familiar study strategies even when their use proved ineffective. Brown et al. (1982) concurred that students resist changing their comfortable strategies for more effective means of studying.

Effective study strategies require active participation from the student. That is, the student needs to attend, interact, reconstruct and elaborate upon ideas, generated through discussion or read from a text, to enhance learning. Developmental courses designed to facilitate students' learning often focus on rules and strategies of study. However, inefficient use of these rules and strategies would still result in inadequate performance by students (Brown et al., 1981).

The purpose of this study was to survey students' perceptions of their study strategies to ascertain whether they see themselves concentrating on rote memory techniques or strategies developed to help them learn how to learn.

Method

All students who were nearing completion of a course in developmental reading for college students were asked to complete the survey instrument. Approximately 20% (N=29) of the students agreed to participate. One group of participants, 13 students from the Developmental Year Program, had been identified as needing assistance in pursuing their academic goals. The remaining 16 students, though not officially identified as academically deficient, chose to enroll in a developmental reading course.

The survey completed by the subjects consisted of 12 open-ended questions concerning six study situations. The open-ended items were used to elicit students' self-descriptions of their behavior regarding topics related to study habits and strategies. Statistical analysis was not attempted due to the small number of participants.

Results and Discussion

The first set of questions on the survey concerned the students' study environment. Answers were given by both groups of students were similar: 53% - Developmental Year (DY); and 56% - Non Developmental Year (NDY) chose residence areas (dormitory rooms or apartments). These were preferred because they were quiet, comfortable, and convenient to materials needed for the study session. The library was listed by fewer than 15% of the students in either group as the place they study.

The times students chose to study were concentrated around the early and late evenings (53% - DY; 88% - NDY). Few of the NDY students gave any set hours for studying, while several of the DY students listed specific study times, including times between classes in the morning and afternoon. There were a few in each group who studied whenever they "got the notion" or "when everything else was done."

Students listed concentration (30% - DY; 31% - NDY) most often as the focus of their study sessions. NDY students also listed "understand what I study" (25%)and "memorize as much as possible" (18%) fairly frequently. The only response other than concentration given by more than one DY student was "complete assigned work" (15%). Descriptions of typical study sessions varied widely but some commonalities did exist. Most NDY students (70%) listed specific study behaviors, such as read notes or text, look for main ideas and examples, highlight text information and make study questions. Some DY students listed behaviors indicative of preparation for study. These behaviors included clear desk, turn on lamp, get comfortable. Others gave answers that could be classified as general advice or vague study behaviors, such as: study one subject at a time, go over material covered in class, review thoroughly, study hardest course first (70%).

On the items concerning reading rate, a majority of students

(60% - DY; 70% - NDY) indicated some degree of flexibility, adjusting their rate according to the difficulty of the material or their purpose for reading. It is interesting to note that DY students were more specific in describing their study behaviors when asked about their rate in relation to the task of reading. For example, one student skimmed the chapter, read more slowly to highlight important ideas, then recited the highlighted information.

The responses to items on lecture class behaviors involved a combination of listening and notetaking (70% - DY; 70% - NDY). Generally, the students in both groups focused their notetaking on repeated information, main points, and examples put on the chalkboard since that information would likely appear on a test. A few students tried to rely totally on memory (less than 10% in either group). Only one student, DY with impaired hearing, mentioned comparing notes with other classmates. That same student and one other (NDY) included "ask questions" in a list to describe their classroom behavior.

When asked about time given to test preparation, "one week prior to test" was listed most often (38% - DY; 50% - NDY) followed closely by "2 to 3 days before the test" (38% - DY; 44% - NDY). Only one NDY student gave a conditional response -- "depending on course." Some DY students (23%) waited until the day or night before the test to begin preparing for it. The strategies listed for test preparation tended to be more specific than those listed for a typical study session. These included self-quizzes, comparing lecture and text notes, rereading, organizing information (outlines study cards), rote memorization, as well as most of the strategies listed for study sessions. It seems that information, assigned readings, or lecture notes become more important if it has to be remembered or used in a testing situation.

The tendency for studying in isolation was noticeable on most of the response forms (90%). Only two of the students mentioned studying with a friend: one for company to keep on task; and the other to share notes, ask questions and discuss information. It seemed these students were wary of group study situations. Subsequent conversations with these students yielded two general reasons for this wariness. One was the fear of being taken advantage of by students who do not contribute to the group efforts, but who expect to receive information and assistance from other group members. The second reason revealed that some students felt studying with others, comparing or commenting on assignments, was tantamount to cheating. To them this apparently meant they were not learning on their own or taking the responsibility for their own learning.

Are students perceiving learning, and studying, as being a competitive or independent task only? Cooperative study situations (Johnson & Johnson, 1975) provide yet another strategy that students can employ in an effort to discuss, apply, and synthesize the information they are eager to learn.

The students' awareness of study strategies and techniques is evident. This type of survey with open-ended questions makes it a bit more difficult for a student to give the "expected" answer that might be recognized in multiple choice items or behavior statements they are to rate. The responses of these students, in both groups, can be interpreted to reflect effective study strategies. But these responses generate other questions. Why do students use particular strategies? Have they made conscious decisions based on how well these strategies work for them? Future research will focus on these questions.

REFERENCES

Brown, A. L., Campione, J. C., & Day, J. D. (1981). Learning to learn: On training students to learn from texts. Educational Researcher, 10, 14-21.

Johnson, D. W., & Johnson, R. T. (1975). Learning together and alone. Englewood Cliffs, NJ: Prentice-Hall.

Simpson, M. L. (1983). Recent research on independent learning strategies: Implications for developmental education. Forum for Reading, 15, 22-28.

Tierney, R. J., LaZansky, J., & Schallert, D. (1982). Secondary students' use of social studies and biology texts. Report prepared in conjunction with research grant from the National Institute of Education.

EVALUATION OF LEARNING DISABLED COLLEGE STUDENTS: A PROFILE ANALYSIS APPROACH

KAY NORLANDER
ANTHONY PAOLITTO
ANIA CZAJKOWSKI
The University of Connecticut

Theoretical Framework

Based upon mandates of Section 504 of the Federal Rehabilitation Act of 1973 (Federal Register, 1977), which stipulates that ". . . no qualified handicapped person shall, on the basis of handicap, be excluded from participation in, be denied the benefits of, or otherwise be subjected to discrimination under any program or activity which receives or benefits from Federal assistance" (p. 2678), many learning disabled (L.D.) students and their advocacy groups are seeking full implementation of this statement in postsecondary settings (Gerber, 1981). The need to provide appropriate support services to this population of students has, therefore, become critical. This need is underscored by several studies designed to describe the current educational aims of this population (White, Alley, Deshler, Schumaker, Warner, & Clark, 1982; Fielding, 1981; Mangrum & Strichart, 1984) with the 1982 survey by White et al., indicating that 67% of young adults diagnosed as L.D. while in public school had plans for postsecondary education.

A portion of Section 504 of the Rehabilitation Act speaks directly to colleges and universities. There can be no discrimination in the recruitment, admissions, or treatment of handicapped students solely on the basis of their handicap. Colleges and universities must make reasonable modifications in academic requirements, where necessary, to fully ensure that the educational needs of handicapped students are being met. While most college or university faculty are acutely aware of their responsibilities toward the "obviously" handicapped student, i.e., the blind or those confined to wheelchairs, these same faculty members are often less sensitive toward students with more subtle "hidden" handicaps. The learning disabled are among those with less obvious handicaps, and are rapidly becoming the largest segment of our handicapped student population.

The problem for postsecondary schools with respect to the learning disabled is twofold. We must make reasonable accommodations for those learning disabled students who have been identified as such and we must also be willing to identify those students who, prior to postsecondary careers, were not identified. In spite of an increased sophistication on the part of secondary school personnel in recognizing and programming for L.D. individuals, a small number of students continue to "slip by." The rigorous demands of college level coursework highlight the deficit areas these two groups of students bring to the classroom setting. Consequently, without individually selected program modification and support, these bright young people are often unable to compete academically. Fielding (1981) claims that there are more than 16 million adults with learning disabilities, characterizing this group as the largest pool of undereducated persons with high potential in our country today.

If postsecondary institutions are to be successful in meeting these new challenges, they must first be prepared to conduct student evaluations which will provide data useful not only for identification purposes, but more critically for the planning of appropriate support services and direct intervention strategies. This paper will specifically address the issues surrounding the evaluation of L.D. secondary and postsecondary students.

Objectives

Objectives for this study are threefold:

1. To select diagnostic procedures and evaluation instruments which will provide valid an reliable information for this population of students.

2. To employ a "profile analysis" technique in the summary of individual student's learning strengths and weaknesses.

3. To conduct a preliminary analysis of group characteristics.

Data Source

The University of Connecticut's (UConn) Program for Learning Disabled College Students provides services to both L.D. students and the faculty who must instruct them. The program is one designed to complement and supplement existing campus services as well as provide direct support services to these students. A psychoeducational evaluation is typically requested or completed for each student. Permission to conduct this evaluation, to be used for program planning and research purposes, is sought from each student.

Methods

Individual Student Assessment

Critical to the provision of appropriate support services to L.D. students is the comprehensive delineation of each student's strengths and weaknesses. Mellard and Deshler (1984) note that college personnel working with L.D. programs must address the following questions in order to ensure effective evaluation of L.D. college students and thereby develop appropriate programs for this population:

> (a) What attributional features are most appropriate for describing the condition of L.D. in post-secondary settings?
> (b) What type of evaluation system is most appropriate to insure monitoring of the identification system to determine its effectiveness in selecting true L.D. students/clients?
> (c) How should intervention models be structured in light of salient identification variables?
> (d) How can identification decision making be structured to insure consistency and fairness across different post-secondary settings? (pp. 8-9)

In response to these questions, a three-step process for identification and program planning is employed by the UConn Program for L. D. College Students. These steps are illustrated in Figure 1 and include screening, diagnosis, and program planning.

Initial screening of "potential" learning disabled college students includes the completion of referral information and an intake interview. Referral forms completed by each prospective student contain information regarding: reasons for referral, description of the problem, previous services, and anticipated or desired services. This referral is reviewed and forms the basis for the initial intake interview. If the student has been previously diagnosed but lacks current evaluative data or if the possibility exists that undetermined learning deficits are interfering with a student's achievement, further assessment is often recommended. Assessment in the areas of general ability or intellectual potential, information processing, reading, writing, spelling, and mathematics as well as study techniques and social/emotional status are, in varying combinations, included for each student. While formal evaluation instruments are employed (see Figure 2 for a partial listing), diagnostic/prescriptive teaching sessions are instrumental in the evaluation process.

Figure 1

Steps to Assessment and Program Planning

OVERVIEW OF ASSESSMENT AND PLANNING PROCEDURES

STEP I: SCREENING

- A. Collection of Referral Information
- B. Initial Interview(s) with Program Coordinators
- C. Meeting with a Learning Specialist

STEP II: DIAGNOSIS

The diagnostic process has two main purposes, to determine eligibility and to identify learning strengths and weaknesses for program planning.

- A. Evaluation Process:
  1. Formal evaluation procedures
  2. Diagnostic/prescriptive teaching
- B. Report Writing Process:

  A team report (Psychoeducational Evaluation) is written, taking into account both formal evaluation and teaching results. The student, as a critical team member, is actively involved in this process.

STEP III: PROGRAM PLANNING

- A. University Planning Meeting
  1. Review evaluation data
  2. Determine program eligibility
  3. Recommend support services
  4. Recommend specific educational goals
- B. Develop Individual Educational Plans
  1. Develop instructional goals and objectives
  2. Initiate a contact between students and program

Figure 2

Evaluation Instruments

STANDARD DIAGNOSTIC BATTERY

- Wechsler Adult Intelligence Scale - Revised (WAIS-R)
- Bloomer Learning Test (BLT)
- Stanford Achievement Test (TASK)

INDIVIDUAL DIAGNOSTIC BATTERY*

- Stanford Diagnostic Reading Test
- Stanford Diagnostic Math Test
- Test of Written Language (TOWL)
- Lincoln Spelling Test
- Gallistel-Ellis Test of Coding Skills
- Diagnostic Reading Scales (Spache)
- Raven's Progressive Matrices
- Reading Skills Diagnostic Test (Level 4)
- Occ-U-Sort (Guide to Occupations)

* An individual diagnostic battery is designed for each student; the above listing represents the instrument most often employed. Additional assessment instruments are selected based upon individual needs.

Group Profile Analysis

Following the examination of individual assessment profiles, an investigation of similarities and differences among 16 cases was undertaken. Cases with the most complete data sets were selected and individual profiles were synthesized. Data from three evaluation instruments were employed in this portion of the study. These evaluation instruments were: the Wechsler Adult Intelligence Scale - Revised (WAIS-R) (Wechsler, 1974), the Bloomer Learning Test (BLT) (Bloomer, 1978), and the Stanford Test of Academic Skills (TASK) (Gardner, Callis, Merwin, & Rudman, 1982). These particular instruments were selected for their usefulness in tapping the areas of ability (aptitude), information processing (learning), and academic achievement (reading, spelling, and mathematics).

Results and Discussion

Individual Student Assessment (Profile Analysis)

During the evaluative period, results are synthesized using a "profile analysis" technique (Cawley, 1977). Patterns of errors across tests and areas of performance are analyzed. In all cases two or more evaluators are involved in the assessment and report writing process. Table 1 represents a psychoeducational profile for an individual learning disabled college student. Translating the evaluative data into understandable and useful information is quite obviously the ultimate objective of this type of assessment process. A format for writing psychoeducational evaluation reports is provided in Figure 3. Thus far, these reports have proved to be extremely useful to both the L.D. student and to the learning disabilities specialist acting in a service delivery capacity. Finally, the findings form a basis for developing instructional goals and objectives for each students.

Figure 3

Psychoeducational Evaluation Report Format

I. Identification Information

II. Background Information

III. Interview Summary/Course of Study

IV. Tests Administered

V. Test Results

   A. Scores
   B. Nature of Results (Strengths & Weaknesses)

VI. Analysis of Results

VII. Statement of Problems

VIII. Recommendations

   A. General Programming Recommendations
   B. Specific Educational Recommendations

Table 1

Psychoeducational Test Profile

Learner's Name: Pam
CA 19
IQ (Full scale) 122
Verbal 122
Performance 115

	CA	APTITUDE	INFORMATION PROCESSING	ACADEMIC SKILLS	STUDY SKILLS	
STRENGTHS	-19-			Read & I.D. Vocabulary		
	-16-			Reading Rate		
			Concept Recognition	Locating Information Decode familiar & unfamiliar words	Motivation	
				Comprehension		
	-13-	Acquired Knowledge	Problem Solving Concept Production Paired Associate Learning			
			Visual STM (Short Term Memory)	English Grammar Math	Time Management Organization	
$\overline{x}$ = 10						
WEAKNESSES	- 7-	Rote Learning	Association Serial Learning	Discriminating individual vowel sounds	Learning Strategies: No verbal rehearsal strategies utilized	
				Recall of written material	Self-questioning limited	
	- 4-		Auditory STM		Tries to recall written material visually	
			Simultaneous STM			
				Spelling common, but complex sight words		
				Discriminating medial vowels or vowel combinations		
				Writing		
				Spelling		

Learner's Name Jeff
CA 30
IQ (Full scale) 129
Verbal 147
Performance 102

	CA	APTITUDE	INFORMATION PROCESSING	ACADEMIC SKILLS	STUDY SKILLS	
STRENGTHS	-19-	Acquired Knowledge		Reading Vocabulary		
	-16-			Reading Comprehension		
	-13-		Association		Motivation	
		Short Term Spatial	Problem Solving Visual STM Auditory STM	English Grammar Math	Time Management Organization	
$\bar{x}$ = 10						
WEAKNESSES	- 7-	Non-verbal reasoning	Serial Learning Concept recog. & production Paired Assoc. Learning	Writing Spelling	Self-questioning while studying Textbook Comprehension tries to learn everything	
	- 4-		Simultaneous STM		"Blocks" information	
	- 1-					

Group Profile Analysis

From this analysis, two distinct group profiles appear to have emerged. Scores within the two groups were found to be remarkably consistent across the above mentioned assessment devices. For the purpose of further discussion, these groups will be referred to as the "superior L.D. group:" and the "average L.D. group." Groups labels were selected based upon the range of WAIS-R IQ scores of each group.

"Superior" L.D. Group IQ Range:	123-132
"Average" L.D. Group IQ Range:	95-118

On a fairly cursory level, a comparison between full scale intelligence scores, as measured by the WAIS-R, and full learning IQ scores, as measured by the BLT, indicated clear differences between the two groups. An examination of IQ differences between the Superior and Average groups shows a 20.9 difference in Full Scale WAIS-R IQs and only a 67.3 difference in Full Learning BLT IQs. These results are presented in Table 2.

In attempting to clarify these findings, the reader must be somewhat knowledgeable with respect to the theories and models of intelligence and learning. A thorough discussion of these theories can be found in the manual of the Bloomer Learning Test (Bloomer, 1984). For the purposes of this discussion it is sufficient to state that the major difference in the WAIS-R and BLT intelligence scores is one of knowledge or information processing. The BLT is a measure designed specifically to examine the manner in which content free or new information is "handled" or learned and thereby infer a student's learning ability. Therefore, the BLT, as opposed to the WAIS-R, does not concern itself with previously acquired information (knowledge), but how students go about assimilating or processing new information.

As the most notable difference between the Superior and Average groups is the discrepancy between "what" they knew and "how" they learned, further examination of specific learning strategies (or lack thereof) is indicated. The Superior Group, while demonstrating a large wealth of acquired knowledge, appears to possess only average information processing skills. Yet the Average Group is near the mean in both acquired knowledge and information processing.

One possible explanation for these findings is that the Superior Group, demonstrating a strength in acquired knowledge, has been able to "survive" without attending to the "mechanics of learning. Additionally, this vast amount of acquired knowledge often seems to interfere with or become counterproductive to systematically learning new information.

Table 2

Mean WAIS-R and BLT IQ Scores

	WAIS-R I.Q.'s[*]			BLT I.Q.'s[*]		
L.D. Groups	Verbal I.Q.	Performance I.Q.	Full Scale I.Q.	Simple Learning I.Q.	Problem Solving I.Q.	Full Learning I.Q.
Superior	126.5 (N=8)	120.3 (N=8)	128.0 (N=8)	100.1 (N=8)	112.2 (N=6)	106.3 (N=6)
Average	109.5 (N=8)	103.1 (N=8)	107.1 (N=8)	95.3 (N=6)	105.3 (N=5)	100.0 (N=5)
I.Q. Difference	17.1	17.2	20.9	4.8	6.9	6.3

[*] All I.Q. Scores have a mean of 100 and a standard deviation of 15.

While obvious differences were found between the learning and aptitude profiles of these two groups, similarities were evidenced on the subtests of the TASK, a measure of academic achievement in a number of subject or content areas (Reading Vocabulary, Reading Comprehension, Mathematics, Spelling, and English). Normed scores for this test are based upon a sample of entering college freshmen and are expressed in percentile ranks, stanines, and grade equivalent scores. Percentile ranks were chosen for use in this study. Percentile ranks for both groups on all TASK subtests are found in Table 3.

Of the five scores derived form the TASK, the percentile ranks in spelling were found to be the lowest for both groups. In fact, spelling was the only area in which scores fell below the 50th percentile. It is also interesting to note that the mean percentile rank for the Superior Group (43.6) was slightly lower than the Average Group (49.3). English was the second lowest area as measured by the TASK for both groups. Additional evaluation data, both formal and informal, highlighted written language as a "deficit" area for both groups of L.D. students.

An examination of the Reading Vocabulary and Reading Comprehension subtest scores also generates several interesting speculations. While both groups scored slightly higher on the Vocabulary subtest as compared to the Comprehension subtest percentile rank, differences between the two groups should not be considered discrepant in light of the overall IQ difference between the two groups. In fact, for the Superior Group, percentiles for Vocabulary and Comprehension, 88.4 and 87.7 respectively, were approximately 10 points below their IQ percentile rank of 97, while the Average Group scored closer on both subtests to their IQ percentile rank of 68.

Conclusion

If postsecondary institutions hope to be successful in providing services to learning disabled students, reliable and valid evaluation practices must be formulated. A profile analysis approach was suggested for the integration of individual student assessment data. Clear report writing leading directly to the development of instructional goals and objectives was recommended.

A preliminary analysis of 16 cases indicated specific differences in acquired knowledge and learning strategies and similarities in spelling and related written language difficulties. Results of the group profile analysis suggest that evaluators must look not only at the general ability of the student and what the student knows (academic skills), but at how the student processes or learns new information.

Table 3

Percentile Ranks: Stanford Achievement Test

		Stanford Test of Academic Skills				
Subject		Reading Vocab.	Reading Comp.	Math	Spelling	English
Superior L.D. Group	1.	98	95	63	31	79
	2.	98	99	99	7	79
	3.	98	99	95	99	95
	4.	98	86	92	38	72
	5.	90	91	75	8	
	6.	90	84	87	64	
	7.					
	8.	47	60		58	56
	$\bar{X}$=	88.4	87.7	85.2	43.6	76.2
Average L.D. Group	1.	36	48	55	34	32
	2.	30	41	83	89	38
	3.	83	60	71	21	56
	4.	83	82	95	2	68
	5.	63	86	99	78	75
	6.	98	77	98	98	86
	7.	89	60	75	34	45
	8.	98	99	58	38	95
	$\bar{X}$ =	72.5	69.1	79.3	49.3	61.9

Blanks indicate test data not available

References

Bloomer, R. H. (1974). Bloomer learning test, manual. Livonia, NY: Brador Publications.

Cawley, J. F. (1977). Meeting of the learning and behavioral needs of handicapped children via a subject matter emphasis: A training institute for general and special educators, diagnosticians and administrators. Portland, ME: Psycho-Educational Associates.

Federal Register. (1977, May 4). 42, 2678.

Fielding, P. M. (1981) A national directory of four year colleges, two year colleges and post high school training programs for young people with learning disabilities (4th ed.). Tulsa, OK: Partners in Publishing.

Gardner, E., Callis, R., Merwin, J., & Rudman, H. (1982). Stanford test of academic skills, Form E, Level 2. New York: Harcourt Brace Jovanovich, Publishers.

Gerber, P. J. (1981). Learning disabilities and eligibility for vocational rehabilitation services: A chronology of events. Learning Disability Quarterly, 4, 422-426.

Mangrum, C. T., & Strichart, S. S. (1984). College and the learning disabled student. Orlando, FL: Grume & Stratton.

Mellard, D. F., & Deshler, D. D. (1984). Modeling the condition of learning disabilities on postsecondary populations. Educational Psychologist, 19(3), 188-197.

Wechsler, D. (1974). Wechsler intelligence scale for children. New York: Psychological Corporation.

White, W. J., et al. (1982). Are there learning disabilities after high school? Exceptional Children, 49, 273-274.

A COMPARISON OF AUDIO-LINGUAL-HAPTIC AND INTERACTIVE VOCABULARY APPROACHES FOR COMPREHENSION TRANSFER OF SECONDARY ESL STUDENTS

VIRGINIA LAUDERMILK AND SHARON W. KOSSACK
Florida International University
Bay Vista Campus

INTRODUCTION

Educators of non-English-speaking students are continually searching for effective methods of teaching English as a Second Language (ESL). The existing methods of accomplishing this task are nearly as diverse as the principles involved. They range from bilingual curricula, with gradual insertion of English education, to immersion--a total English curriculum taught entirely in the second language.

Although ESL teaching and learning problems are common to all instructional levels, secondary teachers are faced with a particularly difficult task. In addition to functional language acquisition students are expected to progress to a level of competency that would allow them to succeed in content courses characterized by heavily conceptual reading and complex language.

In addition to explanatory lecture, secondary teachers tend to use extensive writing practice and reading comprehension exercises to help these students acquire the content material they need in the "regular" curriculum. These same instructors continue to find the second language acquisition process to be extremely slow in relation to academic needs of students.

METHOD

Subjects

The audio-lingual-haptic (control) group for this study consisted of a class of eighteen functionally bilingual students, ranging from fifteen to eighteen years of age. All were classified as B students, that is, they had no less than one and no more than three consecutive years of intensive English training in American schools. In addition to two hours of intensive English instruction, they carried a mathematics class, a tutorial (a study hall, originally designed for remedial science or social studies sessions to prepare students for the "mainstream"), physical education and one elective.

The interactive language development (experimental) group in this study was comprised of twenty-three students, ranging in age from fourteen to eighteen years. All were classified as A students, as they had been in the United States less than six months, had no previous English education and were attending an American school for the first time. At the time of this study, none of the students had been in American schools more than three months. All were scheduled in a three-hour intensive English program in addition to a mathematics class, physical education and one elective.

It is important to note that mathematics and English were the only content area core courses carried by these lower level LEP students.

ASSESSMENT

The post test was a Cloze precedure, created from a textbook passage of average readability, in a content area separate from that of the study, and common to both groups. The control used a text with an average readability on the ninth grade level. The experimental used a text with an average readability of seventh grade. (A pretest was deemed futile because of the obvious disparity in language proficiency between the two groups.)

The Cloze procedure is developed from a passage in which words are regularly deleted. The student must complete the passage by providing the exact words that have been deleted. He relies upon general language knowledge, all levels and types of vocabulary and most heavily, context clues. Thus, the Cloze assesses deep language structure, and the student is classified as potentially frustrated (40%), instructional (40-60%) or independent (60%) in relation to the textbook from which the passage is taken. (Taylor, 1953).

PROCEDURE

This experimental study was conducted over a period of three months, with a post-test only, comparison group design. The control received traditional audio-lingual-haptic instructions, in which students heard the teacher present language structures, repeated them orally and practiced assignments. The textbook and workbook exercises were specifically designed for the learners of a second language. This language acquisition instruction was conducted for half the period. (The regular language arts content was presented during the second half of the period, as these students were scheduled for transition into the mainstream within the year.)

The experimental group received interactive language development instruction in which students used language from the context of teacher-student and student-student conversations that centered on the subject area materials and terms. Due to the variety of first languages represented in the class, students were paired to give an "instructional illusion" of a class half its actual size. Special attention was given to pairing students with different native languages so they were forced to use the only common language for communication: English. Vocabulary development was the key focus. New vocabulary was introduced by the instructor, meanings were derived from conversational context, using appropriate and inappropriate illustrative examples, mnemonics and dictionary entries (as a last resort). Students created vocabulary activities that stressed functional usage, for example: flashcards with Cloze-style sentences, crossword puzzles, matching games with mnemonics, synonyms, antonyms, categorizing devices, analogies, multiple choice-style "computer games", etc. These were then used alone for practice, or paired with other games such as chess, checkers, playing card games, Boggle, etc. Pairs of students were taught different sets of three or four words, and then the pairs were recombined to teach other students their items. Using this technique of peer tutoring, the instructor was able to develop the subject area content as the basis and substance of second language acquisition.

RESULTS

Following three months of instruction, both groups were tested with Cloze tests, constructed from selections of two American History texts on the seventh and ninth grade levels. Scores were analyzed with ANOVA.

Cloze score performances of the experimental group were significantly higher (at the .001 level) than those of the control group, despite the brevity of the study and the considerably lesser initial English language competency of the experimental group (see Table I).

Table I. Means, standard Deviations and t Scores for Cloze Performances of Experimental (Interactive Approach) and Control (Audio-Lingual-Haptic Approach) Groups on two passages used as a Post test.

SEVENTH GRADE PASSAGE

	Experimental	Control
Number	17	13
Mean	44.06	29.8
Standard Deviation	7.83	9.6

NINTH GRADE PASSAGE

	Experimental	Control
Number	11	15
Mean	37.47	9.01
Standard Deviation	2.98	6.8

$t = 14.81 \quad p < .0001$

Discussion and Conclusions

This comparison of comprehension transference of traditionally and interactively educated LEP students suggested the following:

- Direction translation was not necessary for student learning of content.

The use of "real" language, with emphasis on context clue development, was a workable curriculum base for second language development. Older LEP students arrive in schools with one fully developed language. The very concepts, patterns and techniques they employ in their first language can be used as "bridges" as opposed to serving as barriers.

- The best learning took place by "doing".

As in all new language, (English, French, Algebra, Physics, etc.), comprehensive learning is contingent upon vocabulary application and mastery. Every effort was made to teach vocabulary on the functional level. Words that were used were reinforced in lessons, and those that were taught were used repetitively in a variety of ways. Original verbal practice was employed in nearly every lesson.

- Students of different languages relied on each other for necessary instruction, when teachers and aides were engaged with other students.

In the absence of same language teachers, potentially effective one-on-one instruction was provided by peer students themselves. They were divided into pairs of heterogeneous first languages. Besides aiding classroom management, students taught each other by talking through assignments, arguing, defending, providing alternatives and socializing. Using activities designed to include oral exchange among students, the teacher was free to circulate, provide instruction, diagnose difficulties, provide immediate remediation and clarify vocabulary meanings, and repeat instruction.

- Students constructed teaching aids which were modified to accomodate learning levels.

The author found in these students invaluable resources for time-efficiency. Their teaching vocabulary was simpler, and upon rewriting the grade level materials in their own words, the reading levels were automatically lower, allowing other students to proceed at the pace appropriate for understanding.

The evidence that the interactive students outperformed the audio-lingual-haptic students on the Cloze assessment should not be particularly surprising, for it has long been realized that the use of manipulatives and other practical procedures enhances both learning and retention. For this reason, it is logical that the concept should be applied to the secondary LEP learner. The initial disparity in second language competency between these groups was great, yet the comprehension skills learned in one content area were demonstrated in another by the interactive group. The results obtained by the experimental method after a period of only three months are considered astounding.

ESL is only one of many types of compensatory education. It seems likely that our general, fundamental and basic skills students could benefit from similar techniques and activities in striving to attain competency in the content areas. This supposition generates several questions. Although diverse, the subjects in this study were fairly motivated. Would these techniques be as effective in the skills and basic education classes? Would age, maturity or interest levels be significant factors for consideration in a monolingual group of "low achievers?"

In light of the test results in this initial study, it seems that further studies ought to be conducted to explore these variables. The interactive method was clearly more conducive than the traditional audio-lingual-haptic approach to not only second language comprehension, but also the transfer of comprehension from the language arts to another content area (Social Studies).

REACTION PAPER

Commentary on A COMPARISON OF AUDIO-LINGUAL-HAPTIC AND INTERACTIVE VOCABULARY APPROACHES FOR COMPREHENSION TRANSFER OF SECONDARY ESL STUDENTS

Reactor: LOUISE M. TOMLINSON
University of Georgia

Laudermilk and Kossack (1985) have conducted a study which compares two approaches to vocabulary development for English as a Second Language (ESL) students. The control group received a traditional method of instruction involving an audio-lingual-haptic approach which was found to be less effective than the interactive approach applied to the experimental group. The authors attribute the success of the interactive approach to "the use of manipulatives to enhance both learning and retention." Although the reported findings of the study support a method which seems sound, in itself, further substantiation is needed to provide more specific indications of the extent to which the interactive approach is more effective, the specific reasons why it is more effective with a particular group, and implications for its use with other populations.

The authors have stated that an important limitation of the study is a lack of pre-test measures. Although lack of a pre-test eliminates the possibility of assessment of measurable progress from baseline performance to post-test performance, a more serious limitation seems to be the lack of measurements to establish the range of general ability levels which existed among both groups of subjects. It is possible that the general ability level of control group students was much lower than that of the experimental group or vice-versa. Additionally, the disparity in language proficiency between groups, as referred to by the authors, should be specifically identified as a disparity in aural, oral, or written language proficiency. There needs to be some establishment of the fact that both groups do have approximately the same learning capacities or intelligence levels.

The difference between vocabulary approaches is described by the authors as hearing, saying, and writing teacher-selected language structures from textbook and workbook exercises in the control group, while the experimental group was involved in inter-active language from the context of teacher-student and student-student conversations focusing on subject area materials and terms. It is not clear whether the target vocabulary was the same for both groups, since it seems that much of what was emphasized in the interactive group may have evolved spontaneously. This brings to mind the question of generalizability. It is indicated that there was a significantly higher level of generalization from learning in the interactive group to the content of cloze passages from American History texts which were used as a post-test instrument. The authors do not clearly establish the fact that the target vocabulary for both groups was the same and, therefore, raise the question of whether or not the success of group A was due, in part, to a higher level of generalizability inherent to their vocabulary sets.

The question of frequency also arises in comparison of the approaches. Actual time spent on vocabulary instruction is not clearly established as the same for both groups. Since the format and scope of the workbooks for the audio-lingual-haptic group is not delineated, it is not certain whether instructional time was devoted totally to vocabulary or to various comprehension skills.

Other factors which also warrant consideration are native languages and cultures of the subjects and motivation. Cultural background and native language characteristics may have significant influences upon the preferred learning styles and social interaction patterns of students. In some cultures, the interactive mode is inherent to the social context of learning, as in Au's (1981) study of reading instruction in a Hawaiian classroom. Cultural background and the differences between newly arrived and one-to-three year immigrants may also be an important function of the motivation factor in terms of students' expectations and their success-failure experiences in the assimilation process. Given a control group and an experimental group, both of either new students or one-to-three year students, it is possible that there may be no significant differences between results of these approaches.

The study suggests direction for further research in the comparison of audio-lingual-haptic and interactive vocabulary development. In order to determine more clearly the extent to which either approach is more effective, the specific reasons why it is more effective with a particular group, and the implications for its use with other populations, follow-up studies should attempt to control for: homogeneity of ability levels between groups, uniformity of target vocabulary, equal instructional frequency, and compatibility/incompatibility of cultural orientation and motivation with instructional orientation. (For further studies of this nature, information on the Language Assessment Battery (LAB) can be obtained from the Board of Education of the City of New York, Office of Educational Evaluation. The LAB is a series of tests available in foreign languages and English which might be used in determining general ability level, language skills, and, to some extent, experiential background.)

The authors do lead us to some valuable insights and emphasize the use of a vocabulary development procedure that is likely to be effective with a wide variety of populations for many reasons. The interactive mode of instruction may serve to decrease competitive tensions among students; provide positive reinforcement, through the helping roles that students assume with each other; enhance the metacognitive process, not only for individual competency in comprehension and memory, but also the higher level of metacognitive proficiency which Flavell (1976) describes as necessary to explain a concept to others; and foster intrinsic motivation derived from personally relevant contexts, by providing students with the opportunity to link new information to personal schema of previously acquired knowledge.

REFERENCES

Au, K. H. (1979). Using the experience-text-relationship method with minority children. Reading Teacher, 32, 677-679.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In Resnick, L. B. (Ed.), The Nature of Intelligence. New York: Wiley.

Language Assessment Battery, Board of Education of the City of New York, Office of Educational Evaluation.

READING AND LEARNING: THE LEARNING ABILITIES OF THREE TYPES OF POOR READERS

RICHARD BLOOMER
MICHAEL MARTIN
ELIZABETH LEVESQUE
University of Connecticut

Recent research regarding the origins and causes of reading problems increasingly has focused on the multifactorial nature of the reading process. Reading is being investigated as the acquisition and integration of a variety of dynamic processes. A breakdown or deficiency in any of the component functions may result in an inability to read efficiently (Vernon, 1977). During the 1970's the influence of cognitive psychology became very strong, resulting in significant research into memory, knowledge integration, serial and associational learning, learning set, and other basic cognitive processes underlying the reading process. Jorm (1983) states that our understanding of the causes and cures of reading problems is improving due to increased knowledge of how these cognitive processes underlie the complex skills involved in reading.

Research studies are increasingly noting that children with reading problems show very different types of disabilities. Boder (1971), for example, found three types of reading disabled children: a group with visual perceptual deficits, a group with auditory and language deficits, and a mixed group. Mattis, French, and Rapin (1975) found three subtypes: a group with articulatory and other phonological problems, a group with naming and syntax difficulties, and a group with visual perceptual difficulties. Satz and Morris (1981) review this research with respect to learning disabled children.

The present study was designed to investigate the learning skills abilities of poor readers. It was hypothesized that poor readers are not a homogeneous group and that they would be distinguishable by their learning profiles. The subjects were 525 pupils in grades 4 through 6 from four different public schools in a small city in southeastern Connecticut. The mean IQ from school records based on Comprehensive Tests of Basic Skills of this group was 111. Reading scores on the Gates-MacGinitie Reading Tests (MacGinitie, 1978) were collected from class records. Poor readers were defined as those subjects whose reading scores were at least one grade level below their present school grade. This process yielded a sample size of 168 poor readers. The Bloomer Learning Test (BLT) (Bloomer, 1978) was administered to these students. (A description of the BLT is provided at the end of this article.)

These data were used to predict reading level from the BLT scores; this produced a multiple correlation coefficient of .85 and accounted for 72% of the variance. Stepwise multiple discriminant function analyses were performed to determine how effectively the BLT could discriminate good from poor readers. (See Table 1.) The BLT learning scores correctly predicted 72.99% of the grouped cases. The variables which entered into the discrimination functions are *problem solving*, *response integration*, *visual short-term memory*, *concept recognition*, and *activity*.

Table 1

Discriminant Analysis: Prediction of Readers from BLT Scores

Actual Group	Predicted Group	
	Poor Readers	Average plus Readers
Poor Readers	77.7%	22.3%
Average plus Readers	33.7%	69.3%

Percent of ""Grouped"" Cases Correctly Classified: 72.99%

Canonical r = .443

Variables entered into discriminant function	Mean		Univariate F
	Poor	Average plus	
Problem Solving	17.2728	26.2445	45.28**
Response Integration	15.6352	21.2162	3.49
V.S.T.M.	32.0444	33.3959	3.96*
Concept Recognition	46.7072	43.7160	4.98*
Activity	236.7058	235.8458	0.01

*p = .05
**p = .01

Learning skills data acquired from the poor readers was subjected to Q-factor analyses to determine subgroups of poor readers. These data, as well as the entire sample, were exposed to stepwise multiple discriminant analyses (a) in order to identify the poor readers' learning profile; (b) in the prediction procedure to determine the discriminating variables; and (c) in the exploration of the BLT as a diagnostic tool.

Results of the Q factor analyses reveal three factors or types of poor readers from the sample of 168, distinguishable by their learning profiles. Results of the stepwise multiple discriminant analyses indicate that the BLT learning scores correctly predicted 89% of the cases grouped into the three types of poor readers (see Table 2); 90.8% of cases of Factor I poor readers is grouped correctly; 83.9% of Factor II is grouped correctly; and all (100%) of the Factor III students were classified correctly.

Table 2

Discriminant Analysis: Prediction of Three Types of Poor Readers from BLT Scores

	Predicted Group Poor Readers		
Actual Group	**Type I**	**Type II**	**Type III**
Poor Readers Type I	90.8%	6.9%	2.3%
Poor Readers Type II	3.2%	83.9%	12.9%
Poor Readers Type III	0.0%	0.0%	100.0%
Percent of ""Grouped""Cases Correctly Classified:			89.26%
Canonical r = .703			

Variables entered into discriminant function	Mean			Univariate F 2/118
	Type I	Type II	Type III	
Boredom	-1.1040	22.770	-16.3391	16.49
RespInt	5.8335	41.3684	34.0078	37.18
PA Learning	18.6162	11.9732	6.0016	3.18
Activity	247.5267	214.0041	148.007	7.18
VSTM	32.4450	31.1342	30.0036	0.68
StimComp	16.05	16.26	10.00	0.89
VApp	26.8954	25.4250	26.6753	0.50
Prob Solv	17.9927	15.9425	10.3316	1.25

Table 3 illustrates the differences in Intelligence, Learning Intelligence, and in reading grade level for each

Table 3

Mean Language IQ, Simple Learning IQ, Problem Solving IQ, Reading Grade Level and Age in Months of the Three Types of Poor Readers Compared with Average and Above Average Readers

Factor	N	Language IQ	Simple Learning IQ	Problem Solving IQ	Reading Grade Level	Age In Months
I	102	110	93*	78	3.2	137
II	49	103	73	71	2.9	132
III	17	121	73	72	2.2	120
AVG +	357	111	104	116	5.6	127

* Prorated without simultaneous visual short term memory.

of the three types of reading disabilities compared with the average and above average learners.

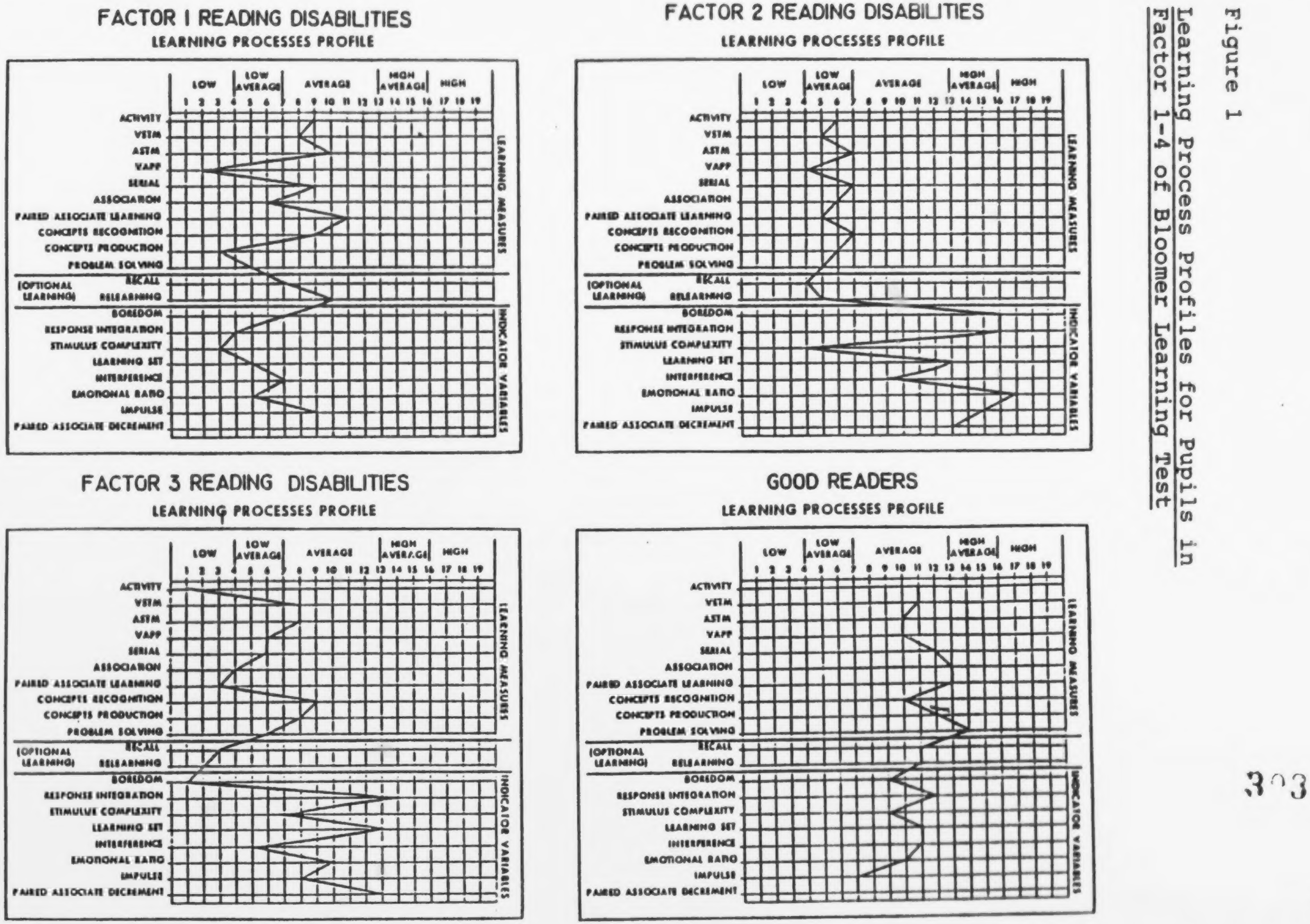
IQ scores taken from the school records primarily represent components of vocabulary and acquired knowledge. All of the groups are average or above in intelligence, meaning they have acquired appropriate vocabulary and factual information at a rate similar to their age mates.

The learning IQ's are developed on a scaling procedure with a mean of 100 and a standard deviation of 15. In each of the three types of reading disabilities, the subjects' Problem Solving IQ, consisting of Association, Paired Associate Learning, Concept Recognition, Concept Production, and Problem Solving, are considerably below their language IQ. Simple Learning IQ, consisting of Activity, Short-term Memory measures, and Serial Learning, was also considerably lower, indicating that each of the types or factors had serious difficulties learning either simple or complex materials.

Learning processes profiles for each of the four factors are presented in Figure 1 and the three involving reading disabilities are described here.

Factor 1 Reading Problems

The Factor 1 poor reader is a child who has adequate ability to learn simple factual material as indicated by average short-term memory for serially presented materials


FACTOR I READING DISABILITIES
LEARNING PROCESSES PROFILE

FACTOR 2 READING DISABILITIES
LEARNING PROCESSES PROFILE

FACTOR 3 READING DISABILITIES
LEARNING PROCESSES PROFILE

GOOD READERS
LEARNING PROCESSES PROFILE

Figure 1

Learning Process Profiles for Pupils in Factor 1-4 of Bloomer Learning Test and adequate paired associate learning abilities. The problem for this child appears to be in handling multiple stimuli, or in what might be termed simultaneous processing. Visual apprehension, simultaneous visual short-term memory, association, concept production, and problem solving all consist of complex stimuli and require simultaneous processing. The Stimulus Complexity and the Response Integration scores indicate a wide discrepancy between the ability to deal with single letters as opposed to words. Thus, this child may learn his letters and phonemes but never blend; or may learn words but have comprehension problems in sentences or paragraphs. The Factor 1 child tends to be emotionally restricted or overcontrolled, which may tend to affect his/her ability to deal with higher order processing. Remedial programs in this instance should be carefully structured to systematically develop higher order processing skills.

Factor 2 Reading Problems

This child is generally a poor learner in all respects. The pattern of learning is nearly flat; the characteristics of this child are seen primarily in the Indicator Scores. Factor 2 children are very easily bored and tend to give up on their work whenever it becomes the least bit difficult or time-consuming. When pressed to finish their work, they will impulsively fill in answers just to complete a page. Further, they are emotionally labile, and usually state that they intensely hate school, teachers, and work. It is clear that we need to focus this child and reduce emotionality before any progress can be made. Generally, this child will need an on-task behavior modification program with impulse control before much success in reading can be achieved.

Factor 3 Reading Problems

This type of reading difficulty is characterized by the very slow, careful way in which the child responds. In contrast to the Factor 2 problem, this child seems to perseverate at a task well beyond the point of profitable learning. This perseveration may have developed because the long term memory of this child is relatively poor. In addition to activity and perseveration, the ability to develop associations are poor for the Factor 3 child; hence, strategies which develop the range of meaning of a word, or which connect and relate ideas or groups of ideas, would appear to be beneficial.

Conclusion

The purpose of this paper is not to develop another typology or nomenclature to which to assign poor readers, but to use these types as illustrative composite cases.

While there are similarities within types, they are intended to illustrate the point that different readers learn differently and that a reading program for remedial or regular class instruction is successful only to the extent that it takes into account the individual propensities to learn.

Success in teaching reading is not dependent upon finding the "best" method, although some methods are clearly superior to others, or the "best combination of methods" as eclectics in the methodology arguments will avow, since many of these "best" practices conflict with one another. The most sensible practice, particularly with remedial readers, consists of determining not only what the child does not know, but also *how he/she learns*, to determine effective remediation techniques.

BLOOMER LEARNING TEST

The Bloomer Learning Test (BLT) is based on a model which attempts to describe academic difficulties on the basis of a pattern of strengths and weaknesses in various learning processes. Poor performance is conceptualized not as the fault of the child, but as being due to the fact that the child has not developed skills in processing information appropriately.

The BLT generates two forms of intelligence score. The first, Simple Learning IQ, might be best termed an input/output IQ. It is the ability of the student to perform rote learning tasks. These are the simple input-output tasks of basic learning and form the content basis for higher order cognitive processes.

Problem Solving IQ, the second type, deals with mediated learning: the interaction between material which is already stored in the memory, the organization of that material, and the input of new material. Problem solving IQ is the ability of the student to perform menial operations on incoming material and, thus, to modify the output to produce relationships or to solve problems.

The two IQ measures yield a Full Learning IQ which represents the child's overall ability to absorb new material and to solve problems.

Profile Scores

The Profile Scale Scores are the chief diagnostic scores produced by the instrument. They are:

ACTIVITY--the rate at which the student responds.

VISUAL SHORT-TERM MEMORY--the amount of successively presented visual material the child is able to hold for processing at one time.

AUDITORY SHORT-TERM MEMORY--the amount of successively presented auditory material the child is able to hold for processing at one time.

VISUAL APPREHENSION SPAN--the amount of simultaneously presented visual material the child is able to hold for processing at one time.

RECALL--the long-term memory ability of the child.

RELEARNING--the effects of past learning on relearning the same material.

FREE ASSOCIATION--the availability of relationships.

PAIRED ASSOCIATE LEARNING--the ability to learn systematic connections between stimuli.

CONCEPT RECOGNITION--the ability to recognize similar characteristics or attributes of stimuli.

CONCEPT PRODUCTION--the ability to identify a concept from exemplars.

PROBLEM SOLVING--the ability to synthesize attributes into a larger whole.

Indicator Scores

In addition to these basic profile scores, the BLT provides seven Indicator Scores which are designed to determine some causes for deviant learning behavior; they describe some of the behavior patterns or habits a student may engage in which are detrimental to learning. These are:

BOREDOM--a measure of frustration tolerance an the tendency to give up to easily.

RESPONSE INTEGRATION--a measure of the ease or fluidity with which a child makes the response.

STIMULUS COMPLEXITY--a measure of the tendency to become confused as the learning requirements become increasingly complicated.

LEARNING SET--the ability to focus and concentrate when required to learn.

EMOTIONAL RATIO--a measure of the degree to which anxiety affects the learning process.

IMPULSE--a measure of the tendency to guess or make random responses.

PAIRED ASSOCIATE DECREMENT--a measure of the relative effectiveness of simple serial learning as opposed to associative learning.

Achievement Expectancy Scores

From the BLT one can derive expectancy scores designed to indicate the level at which a student is presently learning. They represent a standard against which to measure a child's progress, and give a clearer picture of deficits and increments in learning than the use of grade levels as a standard. Four achievement expectancy scores can be calculated:

LEARNING GRADE LEVEL--the general grade level at which the student learns.

READING EXPECTANCY GRADE--the level at which the child's present learning ability predicts he/she should be reading.

MATH EXPECTANCY GRADE--predicts the ability to learn arithmetic computation skills.

SPELLING EXPECTANCY GRADE--predicts the grade level at which, given the child's learning ability, he/she should be able to spell.

References

Bloomer, R. H. (1978). Bloomer learning test. Livonia, NY: Brador Publications.

Boder, E. (1971). Developmental dyslexia: Prevailing diagnostic concepts and a new diagnostic approach. In H. F. Myklebust, (Ed.), Progress in learning disabilities (Vol. 2). New York: Grune & Stratton.

Jorm, A. F. (1983). The psychology of reading and spelling disabilities. London: Routledge & Kegan Paul.

MacGinitie, W. (1978). Gates-MacGinitie reading tests. Chicago: The Riverside Publishing Company.

Mattis, S., French, J. H., & Rapin, I. (1975). Dyslexia in children and young adults: Three independent neuropsychological syndromes. Developmental Medicine and Child Neurology, 17, 150-163.

Satz, P., & Morris, R. (1981). Learning disability subtypes: A review. In F. J. Pirozzolo & M. C. Wittrock, (Eds.), Neurological and cognitive processes in reading. New York: Academic Press.

Vernon, M. D. (1977). Varieties of deficiency in the reading process. Harvard Educational Review, 47, 396-410.

REACTION PAPER

Commentary on READING AND LEARNING: THE LEARNING ABILITIES
OF THREE TYPES OF POOR READERS by Richard
Bloomer, Michael Martin and Elizabeth Levesque

Reactor: RONA F. FLIPPO
Fitchburg State College

The authors of this study focused their conclusions on the very revealing, yet not new, idea that children categorized as poor readers, have wide varieties of reading and learning difficulties and cannot be accurately treated as though they belong in one homogeneous class grouping. The results of this study showed that children who were classified as poor or below grade level readers fall into at least three different combinations of reading and learning problems. This reactor actually suspects that once categorized into these three groups, or any other set of groups, further testing by a reading specialist with more diagnostic, qualitative, and interactive types of reading assessments would yield information regarding the student reading abilities, needs, and interests that would further divide them into many more categories or learning groupings.

Many researchers would focus their results and conclusions on the three factor groups and on the opportunity to distinguish themselves as reseachers who have developed these factor type descriptions. Instead, Bloomer, Martin, and Levesque have used this study as an opportunity to point out how very different "poor" readers can be.

It is important that those of us responsible for training preservice and inservice teachers use this type of data to convince them that ability grouping their students into "high," "average," and "low" groups accomplishes nothing of value toward the learning and reading development of those youngsters. Children in the low reading group have very little in common, regarding reading strengths, reading weaknesses, and reading interests with other children in the low reading group. Ability grouping should be discouraged, and in its place, skill, interest, and learning development groupings should be encouraged. If one or more students have difficulty with short term memory, instruction leading to development in short term memory should be encouraged. However, a need for instruction leading to short term memory development should not be assumed for everyone in a class or every labeled as a "poor" reader.

More studies and more writing need to call attention to the importance of individualization when determining instruction and methods of instruction for individual students no matter how they have been labeled in the past.

THE APPLICATION OF SINGLE-SUBJECT MULTIPLE-BASELINE DESIGNS TO EDUCATION

SARAH D.WEIDLER
State College (PA) Area School District
CONNIE M. MACLAY
The Pennsylvania State University

Between-group designs constitute the experimental methods traditionally applied in the field of education. Often one method of instruction is compared to another, or different conditions of instruction are compared for two groups of subjects. Treatment is sometimes withheld for a control group to permit comparison with an experimental group receiving the treatment.

Another level of analysis, however, is commonly used in clinical psychology but seldom applied in educational settings: single-subject designs. This method of research presents an alternative approach to between-group research. Although there are a number of single-subject designs available, for the purpose of this paper only single-subject multiple-baseline designs will be considered.

The term single-subject design is really an inaccurate label since most researchers apply the design to small groups of subjects. The designs are intended to identify and maximize the uniqueness of the individuals within the groups (Schloss, Sedlak, Elliot, & Smother, 1982). The designs allow a researcher to evaluate the influence of an educational or clinical procedure on a subject's performance by replicating its effect with the same individual over time. Essentially, the subjects function as their own controls.

This type of single-subject design, the multiple-baseline design, is becoming more popular in educational studies. In the multiple-baseline design, measures are made simultaneously on two or more behaviors, subjects, or settings in order to identify a pretreatment (baseline) level. The baseline data describes current level of performance and predicts future performance. If each baseline changes when the intervention is introduced, the effects can be attributed to the intervention. Causality is established by the change in one baseline (where the intervention is introduced) when there is no concurrent change in another baseline (where the intervention has not been introduced) (Kazdin, 1982).

Replication of the results is achieved by applying the treatment to the additional baseline(s). This repeated demonstration that behavior changes in response to the intervention usually makes implausible the influence of extraneous factors (Kazdin, 1982; Tawney & Gast, 1984). The different baselines in the design serve as control conditions to evaluate what changes can be expected without the treatment.

One major advantage of the multiple-baseline design over between-group designs is its flexibility. Individual attention can be given to each client with opportunities to alter and improve the intervention when needed. This is particularly acceptable if the intervention has not been applied in a specific research situation or if it is new or untried (Kazdin, 1982).

There are three common variations of the multiple-baseline design: multiple-baseline across individuals, behaviors, and conditions. In the multiple-baseline across individuals, baseline data are gathered for one specific behavior which is performed by two or more individuals. The baselines refer to the separate individuals within the study. In most designs across subjects, the researcher identifies two or more subjects who demonstrate the same behavior. Baseline observation data are collected on these individuals. The intervention is applied to each subject in turn while data are continuously collected on all subjects. The effect of the intervention is demonstrate when the change in each baseline coincides with the point when the intervention was introduced and not before (Halle, Stoker, & Schloss, 1984; Kazdin, 1982; Tawney & Gast, 1984).

Using a single-subject multiple-baseline design across individuals, Weidler (1985) investigated the effect of cognitive self-instruction on the comprehension monitoring failures of poor readers. The subjects included four elementary-age students all classified as reading disabled. Baseline data consisted of the collection of miscue information from two areas of Goodman and Burke's (1972) Reading Miscue Inventory: Grammatical Relationships and Comprehension. The intervention, cognitive self-instruction, was introduced to one subject at a time while baseline data collection continued for the remaining subjects. It was hypothesized that higher-quality miscues would occur when the intervention was introduced to each subject. Analysis of the graphed data points for several subjects supported the hypothesis since baseline changes wee noted only after the introduction of cognitive self-instruction.

In the multiple-baseline across behaviors, the researcher identifies two or more independent yet similar observable behaviors of an individual or group of individuals. Each target behavior is measured simultaneously and continuously under the same conditions until a sufficiently stable baseline is identified. The intervention (independent variable) is sequentially applied to the target behaviors. This design is especially effective in evaluating an intervention intended to either increase the frequency of an appropriate academic or social behavior or decrease the frequency of an inappropriate academic or social behavior (Tawney & Gast, 1984).

Bornstein, Bellack, & Hersen (1977) reported a study involving four elementary school children who were considered (by their teachers) to be overly passive and unassertive. Baseline data were collected on the ratio of eye contact to speech duration, loudness of speech, and number of requests. The training program (i.e., intervention) focused on each of the three specific skills using a role-playing format with feedback and repeated rehearsal of the correct behavior. The significance of the intervention was established in that each of the baselines changed only after that skill was included in the training.

In the multiple-baseline design across conditions, baseline data are gathered for one specific behavior which is performed by one individual or group (which is considered a unit). The multiple baselines refer to "the different situations, settings, or time periods of the day in which observations are obtained" (Kazdin, 1982, p. 134). The design involves observation of the target behavior in each situation. After a baseline has been established for each situation, the intervention is applied sequentially while continuous measuring of the baseline is carried out across all baselines.

Chiang, Iwata, and Dorsey (1979) report a multiple-baseline study which looked at the disruptive bus-riding behavior of a 10-year-old boy. The baselines were established over two time conditions (i.e., morning ride and afternoon ride). After baselines had been established for both conditions, the intervention was applied to the morning ride (while continuing baselines measures for the afternoon ride). This intervention consisted of a rather elaborate system of feedback and token reinforcement by the driver, teacher, and parents. After a change had been observed during the morning ride, the intervention was applied during the afternoon ride as well. This study provides an example of a multiple-baseline design involving only one subject and only two baselines. Two baselines is generally considered to be the minimum needed for such research.

Table 1 may be helpful in summarizing the three types of multiple-baseline designs. Though these three represent the most common currently found in the literature, combinations and variations are also available.

Table 1

Multiple Baseline Designs

	The researcher must identify two or more different...	The researcher will observe the same...	This design helps to measure...
Across behaviors	behaviors	condition(s) and subject(s)	the effectiveness of the intervention in changing more than one similar yet independent behaviors.
Across conditions	conditions	behavior(s) and subject(s)	the effectiveness of the intervention in changing the target behavior in a variety of situations, settings, and times.
Across subjects	subjects	behavior(s) and condition(s)	the effectiveness of the intervention in changing the target behavior in two or more subjects.

There are several prerequisite steps which may prove useful when implementing a multiple-baseline design (Kazdin, 1982; Halle, Stoker, & Schloss, 1984): (a) the target behavior should be defined, (b) an appropriate recording system should be selected, and (c) the reliability of the measures should be computed. In order to operationally define the target behavior, it must be both observable and measurable. Its definition should meet several criteria: objectivity, clarity, and completeness (Kazdin, 1982). Halle, Stoker, and Schloss (1984) identify six commonly used assessment strategies: frequency counts (e.g., number of words identified correctly), permanent product recording (e.g., number of problems answered correctly), interval recording (e.g., percent of intervals in which an on-task behavior occurred), duration recording (e.g., time spent wearing a hearing aid in school), latency recording (e.g., time needed to identify a word), and momentary time sampling (e.g., percent of observations in which glasses are worn).

Of prime importance is the computation of some sort of interobserver agreement in order to confirm or disconfirm the reliability of the assessment. If frequency counts vary from one observer to another, it would be difficult to ascertain the subject's actual performance. A reliability check between observers also helps to minimize observer biases and lends support for future replication by showing that the target behavior definitions are complete and accurate. Several commonly used methods for interobserver agreement includes frequency ratio, point-by-point agreement ratio, and Pearson product-moment correlation. A more comprehensive explanation of defining target behavior, selecting appropriate behavioral assessment, and calculating interobserver agreement is provided by Kazdin (1982).

Typically, visual inspection is the manner in which any intervention effects are noted for single-subject research. The effects of the intervention are considered clear if there is a noticeable shift in level of the graphed data points when the intervention is applied. The data are plotted graphically so that patterns which might indicate a causal relationship can be inferred (Kazdin, 1982).

Visual inspection depends on characteristics of the graphed data. Usually those characteristics which pertain to changes in magnitude across phases and the rate of such change are analyzed. The characteristics related to magnitude are changes in mean and level, and the two characteristics related to rate are changes in trend and latency of change.

Changes in means across phases refer to shifts in the average rate of performance. If the changes in means across phases are consistent they can serve as a basis for deciding whether or not the intervention was effective.

Changes in level refers to the shift of performance from the end of one phase to the beginning of another. A change in level is independent of the change in mean. Level change can verify what happened right after the introduction of the intervention in order to confirm whether or not the intervention shifted the behavior in the hypothesized direction.

Changes in trend or slope refer to the tendency for the data to show systematic increases or decreases from one phase to another. Typically, a level trend is desirable so that clear treatment effects can be determined. If a baseline shows a trend in the same direction as that expected to result from the intervention, it becomes difficult to evaluate the intervention effects. Multiple-baseline designs are not usually impeded by initial trends in baseline since it is unlikely that all baselines will show a trend in the hypothesized direction and since conclusions about the intervention effects are reached on the basis of the pattern for data across all of the baselines.

Latency of change refers to the period between the introduction of the intervention and changes in performance. The more closely in time that the hypothesized change occurs after the experimental conditions have been altered, the clearer the intervention effects. As a general rule, the shorter the period of time between the introduction of the intervention and the behavior change, the easier it is to infer that the intervention led to change. It would be less likely that other intervening influences may have accounted for the change if the time period between the introduction of the intervention and the behavior change is fairly short.

Variability refers to the fluctuations of the data points as each sample is taken. Ideally, baseline data should show little variability since the greater the variability the more difficult it is to drawn conclusions about the effects of intervention. Small fluctuations are desirable since it describes a more stable and reliable pattern.

Some limitations exist in the multiple-baseline design. In some studies it is quite difficult to identify similar yet independent behaviors or conditions. Interdependence of baselines can confound the results of the treatment (Kazdin, 1982). This can especially be a problem in multiple-baseline designs across subjects since the change in one individual's behavior may influence others who have not yet received the intervention. In such a situation, the relationship of the intervention and the behavior change becomes ambiguous.

Another limitation is the possibility of inconsistent effects of the intervention (Kazdin, 1982). Some behaviors may change when the intervention is introduced while others may not. This raises questions about the generality or strength of the intervention. One should keep in mind that not all baselines need to show the predicted change in order to infer that the intervention was responsible for the change. If several of the baselines show the predicted effect, an exception may not negate the drawing of inferences but may, instead, suggest areas for further study.

A third limitation is the possibility of a prolonged baseline which could lead to methodological problems in that the behaviors may sometimes improve even before the intervention is introduced. Prolonging baselines may also be time-consuming, impractical, distracting, and costly (Tawney & Gast, 1984). There is also (as with the no-treatment control group) the ethical question of withholding treatment. Some variations proposed to avoid this situation include introducing the treatment to two behaviors at the same time or observing at longer intervals rather than daily or hourly.

The advantages of the multiple-baseline designs are:

1. The multiple-baseline designs are an effective means of evaluating educational procedures since they allow continuous monitoring through both learning and retention stages (Tawney & Gast, 1984).

2. The intervention is applied to only one (or two) baselines at a time allowing the researcher to focus more closely on the changes which may occur (Halle, Stoker, & Schloss, 1984).

3. The multiple-baseline designs are not dependent upon large numbers of subjects.

4. The multiple-baseline designs allow the gradual application of the intervention which permits the researcher to test its effectiveness on a small scale and make modifications if necessary (Kazdin, 1982).

5. The multiple-baseline designs do not require the withdrawal of treatment to establish a causal relationships.

Multiple-baseline designs may be appropriate for researching a variety of educational questions. They allow the researcher to look closely at an intervention and the applicability to a variety of behaviors, conditions, and subjects. They are clearly useful tools which are well suited for both clinical and classroom research.

References

Bornstein, M. R., Bellack, A. S., & Hersen, M. (1977). Social skills training for unassertive children: A multiple-baseline analysis. Journal of Applied Behavior Analysis, 10, 183-195.

Chiang, S. J., Iwata, B. A., & Dorsey, M. F. (1979). Elimination of disruptive bus riding behavior via token reinforcement on a "distance-based" schedule. Education and Treatment of Children, 2(2), 101-109.

Goodman, Y. M., & Burke, C. (1972). Reading miscue inventory. New York: Macmillan.

Halle, J. W., Stoker, R., & Schloss, P. J. (1984). Facilitating teacher-conducted research: A tutorial on single-subject design--the multiple-baseline. The Volta Review, 86(2), 89-101.

Kazdin, A. E. (1982). Single-case research designs: Methods for clinical and applied settings. New York: Oxford University Press.

Schloss, P. J., Sedlak, R. A., Elliot, C., & Smothers, M. (1982). Application of changing-criterion design through token reinforcement. Journal of Special Education, 16(3), 359-367.

Tawney, J. W., & Gast, D. (1984). Single-subject research in special education. Columbus, OH: Charles E. Merrill.

Weidler, S. D. (1985). The remediation of disabled readers' metacognitive strategies via cognitive self-instruction. Unpublished doctoral dissertation, The Pennsylvania State University, University Park.

REACTION PAPER

Commentary on THE APPLICATION OF SINGLE-SUBJECT MULTIPLE-BASELINE DESIGNS TO EDUCATION by Sarah Weidler and Connie Maclay

Reactor: CAROLE F. STICE
Tennessee State University

The purpose of this paper was to describe one type of single-subject research design, specifically the multiple-baseline design. According to Weidler and Maclay, the multiple-baseline design measures two or more behaviors, subjects, or settings simultaneously in order to establish a baseline. If the baseline(s) changes when and where the treatment is introduced causality is determined. The paper provides description and examples of the most common variations on the single-subject multiple-baseline design: across individuals, across behaviors, across conditions.

Single-subject (and small groups) multiple-baseline research designs offer an alternative to more traditional between group experimental designs. It also provides an alternative for those researchers who are not completely comfortable with purely qualitative, descriptive, ethnographic kinds of research.

This paper is well written, clear and concise. Both the reference list and the table are helpful for those who wish to know more about single-subject research.

The one note of caution raised during the discussion following Weidler's presentation of the paper is worth mentioning here. Some data must be collected over a sufficiently long time for trends to become clear.

ETHNOGRAPHIC RESEARCH ON JUNIOR GREAT BOOKS: DISCUSSION LEADER BEHAVIOR

NANCY BORAKS
Virginia Commonwealth University

JULIE EARLY
Chesterfield County Schools, Virginia

EILEEN SABLE
Fairfax County Schools, Virginia

The Junior Great Books (JGB) Program is patterned after the adult Great Books program. Both programs expose readers to the classics and other literature which lends itself to interpretive discussions. The JGB program serves students in grades 2-12. Programs are led by discussion leaders who are trained to develop student comprehension; specifically, interpretation of the author's meaning (Dennis and Moldorf, 1983). In JGB children generally preread a story and meet with a trained leader once a week to discuss the story.

The framework of the Great Books program developed by Adler and Hutchins seems to be best articulated by the work of I.A. Richards (1929). There is a focus upon text-based interpretation, testing of one's ideas against anothers and verification of statements -- not through experience but through textual reference.

The JGB program is widely used. According to JGB promotional literature over one million children have been involved in this program. In Virginia the program is often used to provide enrichment for gifted children. Yet, research on the program is limited largely to surveys of leaders and students. Several aspects of the program warrant closer examination.

One critical component of the program is the discussion leader. JGB discussion leaders are trained in a two day workshop. All who apply and pay the required fee can attend the workshop. All who attend for the requisite time are given leader certificates; no posttest is given at the conclusion of the workshop. There are no systematic follow-up observations on leader impact. In a 1984 pilot study, variation in leader behavior was noted. Both student and leader ability seemed to impact upon the quality of discussion.

BACKGROUND OF THE STUDY

Goal: The goal of this study, part of a long term analysis of the JGB fifth grade program, was to examine the impact of JGB leader behavior and student ability on a) student response during discussion and b) student acquisition of desired JGB goals.

Procedure

Very little is known about the impact of actual behavior of leaders and students during JGB discussions. All variables in the JGB question/response interaction have not been identified. Thus, an experimental study is still inapporpriate (Guba, 1980). An ethnographic approach was used (Guba, 1980). Four JGB discussion groups were systematically observed and taped over a one year period. Four to eight observations of each group were made. Observers took verbatim notes and/or taped the sessions as unobtrusively as possible. Transcripts were analyzed for question/response behavior. A final brief questionaire was administered to the leaders and a random sample of JGB students.

To provide comparative data and to get some estimate of the impact of students' ability on final behavior in relation to JGB goals, pre and posttests were given to identified gifted and identified nongifted students who were and were not enrolled in JGB programs. Only an analysis of posttest data is given here.

Leader Participants

All leaders were female. All had taken JGB training between 1982-4. Of the four leaders observed systematically, three were or had been teachers - one was a parent. Four additional JGB leaders (three teachers, one librarian) were observed once and/or interviewed. Student data from these groups was not collected.

Student Participants

Eighty-six fifth grade students were involved in the study. These children were in intact groups. Both rural and urban area schools were involved. Intelligence tests scores were used to designate children as gifted.

Leader and Student Questionnaires

Leaders were questioned to determine what they believed children gained from JGB and what they/the program did to make this possible. Ten students (5 girls, 5 boys) randomly selected for interviews, were asked what they gained from the JGB program and what they liked/disliked about the program.

Student Test

The student test was developed to tap JGB goals and objectives indicated by JGB training materials and student exercises. There seems to be four goals in the fifth grade program: 1) deeper appreciation of literature, 2) improved interpretive comprehension 3) ability to generate high level questions, and 4) understanding of the goals of a literary discussion. Both open ended and forced choice items were used to assess these JGB goals.

Rater reliability was estimated using a weighted kappa (Fleiss and Cohen, 1973). Five education students coded a random selection (5) of the student questionnaires. Item reliabilities ranged from .36-.92.

Limitations

This study is considered exploratory because 1) intact groups are involved 2) the total number of groups is small and not considered representative and 3) the student test instrument used is in a developmental stage. Moreover, fifth grade JGB program materials have changed somewhat in the past year.

Tentative conclusions are drawn where observations, JGB leader and student comments, and student behavior on the final test are congruent.

RESULTS OF THE STUDY

Three facets of JGB leader behavior are described: 1) the role of the leader, 2) questions used in discussion and 3) follow-up questions. As each facet of the program is discussed, observed leader behavior and JGB guidelines are contrasted with guidelines recommended in recent research. In addition, the impact of leader behavior on student behavior/achievement is analyzed. Finally, suggestions for related JGB program changes and needed research are made.

I. ROLE OF DISCUSSION LEADER

There is a large body of research on the effective behaviors of discussion leaders. This research suggests that the discussion leader is to promote discussion by 1) crystallizing contributions and keeping participants on topic 2) raising relevant questions, 3) clarifying the goals of the group 4) establishing discussion procedures 5) making helpful contributions and 6) making supportive statements (Hyman, 1980).

The stress in JGB leadership training has been on the first two behaviors. The latter two are discouraged. Prior to 1984 JGB leaders were not encouraged to discuss group procedures but the 1984 leader guide encourages leaders to 'go over the rules of discussion' at the start of each session. During discussion, the JGB leader is to ask questions - not suggest answers; to get students to think for themselves by promoting responses - not directing responses; and to keep the discussion focused.

Observed Leader Behavior

The actual behavior of leaders during JGB discussion conformed more to guidelines suggested by recent research than to JGB guidelines. In several instances specific JGB guidelines were broken. The following behaviors were observed:

1. Leaders do reinforce student responses.

Through nonverbal and verbal language, leaders often praised or reacted negatively to student responses. This behavior seemed to be used when the child's answer was promoting a negative peer reaction, or when the child said something the leader felt of value. In some cases reinforcement seemed a natural result of leader interest in the issue:

Leader: Were the kids happy when it stopped raining?
Sue: I think that after the rain they might have been happy again...might have some of the happiness left in them
Leader: (thoughtfully) Yes, but it would have to last years..

In other cases, the leader habitually used reinforcing comments like "good", "that is interesting". Leaders' use of reinforcing statements did not seem to inhibit children's reactions, for example:

Leader: OK, any ideas on that, Laura?
Laura: Only because the kids teased her
Leader: OK, that's good.
Jeb: (persisting in defending a prior response) I don't think they wanted to play with her..... So like Judy said ......

In all groups there was student reaction to peer responses in the form of nonverbal and verbal reactions, e.g. nodding at, or groaning at answers. In most instances, student reactions to peers came in the desired JGB form of agreeing or disagreeing with peer responses.

2. Leaders introduce outside information.

The introduction of outside information in JGB should be in the form of a question. If the teacher wants to point out that the children may have been excited by the sun, she would say "Because the children were excited by seeing the sun, is it possible...?"

However, no leader seemed adept at using this form. Several leaders introduced information directly. In the case of one leader this information was usually incorrect.

In most cases, the introduction of information did seem warranted by the great deal of misinformation introduced by students. In some groups and instances peers corrected misinformation. In other groups the information stood unconfirmed, but uncontested.

3. Leaders introduce a variety of alternative activities.

Teachers told of or were observed having students: illustrate selections, look up and learn definitions of given words, act out selections, write answers, summarize the selection, and watch films of works read and compare/contrast film/text versions. Some of these activities worked quite well. Leaders stated that they used these techniques to integrate JGB with other programs (writing); others felt they were crucial to maintaining group interest. One leader who followed "to the letter" JGB discussion guidelines found students rebelling and parents demanding changes. New JGB leader guides list and apparently encourage alternative activities.

4. Leaders allowed nonreaders to join discussion.

Students who had not read the story, according to JGB guidelines should not have been allowed to participate. In actuality several leaders allowed this noting they did not want to send a child back to the class or deal with discipline problems a non-participating child might cause.

5. Leaders do not set group goals or procedures.

No leader was observed presetting the goals or rules of discussion. There were occasions when this might have helped the leader. For example, one group began with the leader asking questions written by student; these were asked in a rapid fire manner - all children would simultaneously give answers. Then, the leader would move onto the next question. In this instance, discussion of peer answers occurred only when the children had disparaging remarks to make about that answer. A presetting of goals and procedures might have acted as a check on this routine.

The leaders do know about JGB goals. Student and leader responses to the final questionnaire indicated that both see sharing ideas, getting ideas from peers, agreeing and disagreeing as JGB program characteristics.

6. Leaders do keep students on topic.

In almost all instances the leaders succeeded in keeping students focused upon the topic, i.e. children spoke to the question at hand. However, in most instances each questioning episode lasted less that 6 minutes and subsequent questions were not related to the initial question. Overall, the result of keeping children on topic was great sidcontinuity in the total discussion.

Impact of leader role

No leader or student expressed concern with most of the behaviors noted above: use of reinforcement, peer positive and negative reinforcement or the disregarding of JGB guidelines. The role of the JGB leader is partially to insure that discussion is shared inquiry and thus children view a leader's remark as a prompt to reflect, and not as an indication of a poor or incorrect answer. Student responses to the questionnaire indicated that fifth grade students will view a teacher's prompts positively, if these students are gifted - and whether they are in JGB or not. When asked why a peer asks them to verify their response, or why a classmate disagrees with their response, most students (gifted, nongifted, JGB, non-JGB) viewed their peers contribution in a positive light. JGB leaders and students do see shared inquiry as a major goal and value of the program. Leaders see their role as promoting expression of and evaluation of different views. In actual practice, a variety of views are not always considered objectively by peers or leaders. In all instances leaders express a need to vary the JGB format.

LEADER USE OF DISCUSSION QUESTIONS

There is some divergence between the type of questions promoted by JGB and the type of questions recommended by research on teacher questions. Currently there is recognition that children who start reading with broad questions tend to have greater comprehension than peers who read for details.

In JGB children are to read once (with no specific purpose) and then reread to note concerns or questions. Guidelines for discussion questions are specific. JGB guidelines state: "General questions make it hard to think just as vague directions make it hard to act". JGB guidelines and training emphasize that the leader is to carefully prepare and organize story questions. In addition, children are trained to write interpretive questions and these can be used in promoting dicussions. Interpretive questions are also provided in the children's text and leader's guide.

Observations revealed that JGB guidelines for leader development of discussion questions are not followed. The following behaviors were observed:

1. Leaders rarely prepare and organize their own questions.

Most leaders use leader guide or children's questions. Leaders stated they felt the leader guide questions were better than their own. Leaders also explained that children wanted the questions they had written used. One child was observed leaving a discussion frustrated because his question was not used. He muttered "It was a darn good question."

Use of children and text questions contributed to the episodic nature of discussions because these questions were not organized into topics and because children tended to base questions on a single segment of text or character behavior not the story theme.

2. Discussions focus upon story detail.

JGB questions are to capture the main elements of stories or important questions of interpretation. Children's questions did not do this, children tended to question character behavior or story details. Moreover, they rarely asked questions that seemed to be points of difficulty-children seemed to ask questions on points they could understand.

Field notes indicate that in many instances the discussion did not insure that children actually understood the story theme. In a final interview one child indicated he was still bothered by the fact he did not understand a story. Leaders are not encouraged to explore larger reasons for potential misinterpretations. Analysis for a child's answer is in relation to small segments of text - not the larger text structure.

Impact of Questioning Behavior

The use of children's questions and focus upon segments of story help with one JGB goal but defeat another. JGB children do learn to write interpretive questions. When JGB children were asked to write five questions for a story, both gifted (61%) and nongifted (71%) asked more interpretation questions when compared to nonJGB gifted (25%) and nongifted (48%) students. Interestingly enough nongifted students not in JGB asked more interpretive questions than their gifted nonJGB peers. This raises questions about the desirability of a heavy emphasis on interpretive questions.

JGB children do not seem to increase literary appreciation. This seems to be an ability existing in gifted children. For example, when students were asked why they liked a particular ending to a story - 50% of JGB and nonJGB gifted list elements of style and structure but less than 20% of nongifted JGB and nongifted nonJGB were influenced by this. Nongifted children tended to consider likeability of behavior/characters in selecting an ending.

Nongifted JGB children, however, seem to make more progress in dealing with interpretive questions. Gifted children in JGB make no progress in this area. Thus, when asked to respond to a question about the author's intent in leaving out some information in a story, JGB nongifted children (35%) did significantly better than their nongifted nonJGB (13%) and better than gifted JGB (26%) and non JGB (25%) peers.

LEADER FOLLOW UP QUESTIONS

JGB guidelines provide helpful steps for insuring that there are follow-up questions. As noted above, prior to meeting with their group, leaders are to organize questions into categories. This would provide a series of questions on a given topic which would allow the leader to have effective follow-up questions, the use of follow-up questions varies. The following behaviors were observed:

1. Children are asked to prove points by referring to text.

Children are encouraged to refer to specific lines of text to prove their point. Often, when children were asked to read a selection to prove their answer, they referred to unrelated passages. When peers did not pick up on this, several leaders would become directive, using phrases as "Don't you think.." and "Wouldn't that mean.." In most instances, leaders had trouble helping children in this way. A typical example:

Leader: Jenny, can you show why you think that?
Jenny: (reads from page)
Leader: What is it you are trying to tell us?
Jenny: I don't know.
Leader: Here's another question.

2. Leaders do use children's prior remarks to generate questions.

Teachers do use standard JGB follow-up questions as "can you find that in the story, what do you mean, why do you say that". However, unless fellow students use this opportunity to clarify why they agree or disagree with peers, these leader statements simply lead into a new line of discussion. Leaders rarely attempted resolution to clarify or organize points made. Occasionally, to conclude the discussion of a question - one leader rephrased what she believed to be the best answer to the question.

3. Leaders do not correct misstatements of fact.

According to JGB guidelines, follow-up questions can lead to self-or-peer correction of misstatements. In one typical instance a child gave three incorrect facts in responding to q question.

The leader tried the question "Do you agree/disagree with that?" Children in the group intent upon making other points carried the discussion in a new direction.

The leader's ability to ask follow-up questions on potentially correct points is excellent. Their ability to deal with incorrect responses is not. As the discussion excerpt below indicates, the leaders' own ability/alertness (or lack of same) influenced the effectiveness of follow-up questions:

Leader: Why did the children say she was lying?
Allison: If they had not seen it or felt it..couldn't..
I don't know how to say it..
Leader: Were they jealous?

In several instances discussions became convoluted as the leader tried using a series of indirect questions to lead students into dealing with misinformation. This approach was rarely successful.

4. Follow-up questions generate a variety of responses and reasons for responses.

Most leaders were adept at using follow-up questions to generate a variety of responses from children. In many instances children were able to verify responses. Strongest verifications came from indirect references to text and references to logic or personal experiences. However, JGB goals of having the group generate as many possible interpretations as possible seems to have a slightly depressing effect on individuals. Thus on the final test, nonJGB children when asked to give as many reasons as possible for a character's behavior, listed more possibilities than JGB children. JGB children may have found that stating one or two possibilities sufficed, leaving room for others to give alternative answers.

Impact of leader follow-up questions

Leaders use follow-up questions effectively when they are not challenged to deal with incorrect or misguided responses. This may be why JGB children who enter the program able to respond to requests to provide proof do so - and these children tend to dominate discussions. Children who do not enter with these abilities do not seem to develop them in JGB. Leader follow-up questions are most successful in clarifying points.

There was concern for the amount of misinterpretation indicated by children's responses and concern about leaving children with that misinformation.

RECOMMENDATIONS:

Leaders know but do not follow all JGB guidelines. In some instances this promotes better discussions. In other instances, failure to use JGB guidelines seem to obviate good discussions. Leader behaviors seem to promote greater changes in the nongifted student. Changes seem needed to a) help leaders with follow-up questions, b) insure better comprehension of total story structure and c) provide greater learning opportunities for the gifted. Results of this study indicate that the following recommendations warrant consideration:

1. Greater variations in the discussion format should be encouraged. Leaders indicate that more variations are required when students do not immediately enjoy or understand a story. JGB might provide more options and encourage those activities which promote awareness of total story structure and theme.

2. There seems to be a strong need for leaders to reiterate at the beginning of each session JGB procedures and discussion goals. Specific guidelines should be given as to how children are to react to differences of opinion AND misstatements.

3. Some screening or monitoring of 'certified' leaders should be instituted. Leaders might tape and evaluate with peers several sessions each year.

4. Guidelines for promoting continuity in discussion should be revised to insure such continuity. For example, at the beginning of each session-leader and students could start by organizing questions and together decide which category of questions would best capture important story elements.

5. JGB text questions should provide a series of questions on story theme. Leaders might be encouraged to start with this question or with a student rendering a story summary. Guidelines for dealing with frequent misinterpretations of the story line should be provided.

6. This use of JGB with gifted children must be reevaluated to determine how discussions can truly provide enrichment. Right now, JGB seems to provide an opportunity for the gifted to display - not increase - abilities.

CONCLUSIONS

There is great variability in JGB leader ability and use of JGB quidelines. In some instances the ability of the leader to profitably use JGB guidelines is a function of the ability of students in her group. Changes seem warranted in leader selection and training. Further research on the impact of Junior Great Books on students is needed.

REFERENCES

Andresen, O. (1969). "The significance of profundity in literary appreciation" Reading Research Quarterly, 5, 100-118.

Dennis, R. and Moldorf, E. (1983). Handbook on interpretive reading and discussion. The Great Books Foundation: Illinois.

Fleiss, J. and Cohen, J. (1973). "The equivalence of weighted kappa and the interclass correlation coefficient as measures of reliability," Education and Psychological Measurement, 33, 613-19.

Guba, E. (1980). "Starting up, gaining entry and getting established", A paper presented at the American Educational Research Association.

Hyman, R.T. (1980). Improving Discussion Leadership, Teachers College Press: New York.

Richards, I.A. (1929). Practical Criticism, London: Trence and Tubner

EVALUATING MATERIALS: DO THEY TEACH PHONICS GENERALIZATIONS?

MARTHA D. COLLINS
Louisiana State University

The major reading material used in the elementary classroom is the basal reader with teachers relying to a great extent upon the teacher's manual to direct their teaching (Ingersoll & Smith, 1984). This study was designed to determine the extent to which phonics generalizations are taught within the five most frequently used current basal readers. Although the study was conceptualized and begun prior to the publication of Becoming a Nation of Readers (1985), the two recommendations of the Commission on Reading regarding the need for systematic phonics instruction in beginning reading lend support for an investigation of this type. Only basal reading materials were evaluated in this study due to their predominance in the classroom--workbooks and supplementary phonics materials are used primarily to supplement the basal.

The procedure followed in evaluating basal materials is discussed below.

(1) The five most frequently used basal readers were identified. To do this the readers used in Sorenson's (1982) study of basal materials were first considered. Secondly, each state department of education within the continental United States was contacted to determine (a) if they had a state textbook adoption or approval list and (b) which basals were on their most recent elementary reading list. From this information the five most frequently listed basals were selected. They were (in alphabetical order):

    Ginn Reading Program, K-8, 1984
    Harcourt Brace Jovanovich Series, 1983
    Houghton Mifflin Series, 1981
    Macmillan Reading Series, 1983
    Scott, Foresman Reading Program, 1981

(2) Teachers manuals for each series, grades one through six were reviewed to determine which phonics generalizations were taught. Rules or generalizations which were included but not directly taught were not considered.

(3) All generalizations which were directly stated and had directions for instruction were listed.

(4) Materials were reviewed a second time along with a copy of the scope and sequence guides published by the companies to ensure that all appropriate generalizations were included.

(5) Twenty experienced elementary teachers and doctoral level graduate students in reading reviewed the list of phonics generalizations to determine the accuracy of the statements.

(6) Those generalizations appearing in at least two of the series were included in the final listing.

RESULTS

The findings of this evaluation of basal reading materials supported the idea that phonics generalizations do appear in current reading materials with procedures for teaching provided. Twenty-eight generalizations were included in two or more of the series. These generalizations are presented in Table 1. Twenty-eight others appeared in only one series. Additional evaluation indicated that there was a general consistency in the sequence of the generalizations with the majority presented in grades one through three. Most series did not begin phonics instruction until the preprimer level and four suggested teaching consonants first. One of the five series introduced vowels first.

None of these series used a synthetic approach for teaching phonics; generalizations were based on an analytic approach. The wording of the generalizations reflected the current terminology of letters or letter patterns representing sounds although most of the series used the terms digraphs, diphthongs, and blends rather than clusters.

Further comparisons were made between the generalizations identified in this study and those identified by Clymer (1963) and Sorenson (1982). Clymer's study included eight basal reader series, Sorenson's study included five, and both evaluated grades one through three. The five series used by Sorenson were the same used in this study but were earlier editions. Clymer selected 45 generalizations examined in further study, although there were about 125 originally identified. Sorenson listed 37 generalizations; however, there were only 16 found in more than one of the basal series. This study identified 28 generalizations appearing in two or more of the series with 28 others appearing in only one of the series. A display of this information is provided in Table 1.

TABLE 1

PHONICS GENERALIZATIONS IDENTIFIED IN THREE STUDIES

Collins (1986)	Sorenson (1982)	Clymer (1963)
	Two vowels together in a 1 syllable word usually stand for the long vowel sound.	When there are 2 vowels side by side, the long sound of the first is heard and the second is usually silent.
When a vowel is between two consonants in a word, the vowel is short. The pattern CVC and CVCC produce the short v sound.		When a vowel is in the middle of a 1 syllable word, the vowel is short.
When e comes at the end of a word, and is the only vowel in the word, the e is long.	When e comes at the end of a word and is the only vowel in the word, it has a long sound.	If the only vowel letter is at the end of a word, the letter usually stands for a long sound.
	In words in which the vowel is followed by a consonant and final e, the long sound of the vowel is usually heard.	When there are 2 vowels, 1 of which is the final e, the first vowel is long, the e is silent.
When the letter r follows a vowel letter, it changes the sound that the vowel letter usually stands for.	When the letter r follows a vowel, it changes the sound the vowel usually stands for.	The r gives the preceding vowel a sound that is neither long nor short.
	Two vowels together often stand for the long vowel sound of the first one.	The first vowel is long the second silent in the digraphs ai, ea, oa, and ui.
		In the phonogram ie, the i is silent and e has a long sound.
		Words having double e usually have the long e sound.

If the letter a stands for the long vowel sound, the letter e at the end of the word is a marker to indicate that in these words the letter a stands for the long vowel sound.		When words end with silent e, the preceding a or i is long.
The letters ai and ay also represent the long a sound.		In ay the y is silent and gives the a its long sound.
		When the letter i is followed by the letters gh, the i usually stands for its long sound and the gh is silent.
		When a follows w in a word, it usually has the long a as in was.
		When e is followed by w, the vowel sound is the same as represented by oo.
When o is followed by a w, the o and the w together sometimes stand for the sound you hear in flow and how.		The 2 letters ow make the long o sound.
		W is sometimes a vowel and follows the vowel digraph rule.
	The letter y usually stands for the long i sound when it appears at the end of a one syllable word.	When y is the final letter in a word, it usually has a long vowel sound.
	When y appears at the end of a word, it is usually a vowel sound with long i or long e sound.	When y is used as a vowel in words, it sometimes has the sound of long i.

		The letter a has the same sound when followed by l, w, and u.
		When a is followed by r and final e, we expect to hear the sound heard in care.
		When c and h are next to each other they make only one sound.
The ch spelling for c may occur at the beginning, middle, or end but tch comes generally at the end and the t in the middle.		Ch is usually pronounced as it is in kitchen, catch and chair, not like sh.
The letter c can stand for 2 sounds. The sound of s as in celery and k as in came.	The letter c usually has the sound of s when followed by e or i.	When c is followed by e or i, the sound of s is likely to be heard.
		When the letter c is followed by o or a, the sound of k is likely to be heard.
When g is followed by e, i, or y it usually stands for the sound heard in gem. When g is followed by a, o, or u it usually stands for the sound heard in gun.	When g comes before e, i, or y in a word, the g stands for the sound of j.	The letter g often has a sound similar to that of j in jump when it precedes the letter i or e.
		When ght is seen in a word the gh is silent.
	The letters kn at the beginning of a word make the n sound only.	When a word begins with kn the k is silent.
	The letters wr at the beginning of a word make the r sound only.	When a word begins with wr the w is silent.

	Double consonants usually stand for a single sound in the middle or at the end of a word.	When two of the same consonants are side by side only one is heard.
Sometimes when the ck sound is in the middle of a word or at the end of a word it represents the sound of k.		When a word ends in ck, it has the same last sound as in look.
		In most two syllable words the first syllable is accented.
		If a, in, re, ex, de, or be is the first syllable in a word, it is usually unaccented.
		In most 2 syllable words that end in a consonant followed by y, the first syllable is accented, the last is unaccented.
		One vowel letter in an accented syllable has its short sound.
		When y or ey is seen in the last syllable that is not accented, the long sound of e is heard.
		When ture is the final syllable in a word, it is unaccented.
		When tion is the final syllable in a word, it is unaccented.
		In many two- and three-syllable words, the final e lengthens the vowel in the last syllable.
When a word has a vowel followed by 2 consonants and another vowel; the syllables are usually divided between the 2 consonants.		If the first vowel sound in a word is followed by 2 consonants, the first syllable usually ends with the first of the 2 consonants.

When a word has a vowel followed by a single consonant letter and followed by another vowel letter, we often divide the word after the first vowel letter. The first vowel letter stands for the long sound.		If the final vowel sound in a word is followed by a single consonant, that consonant usually begins the second syllable.
When the second syllable of a word ends in -le, you divide the 2 syllables between the vowel and the consonant preceding the le.	A word that ends in a consonant followed by the letters le breaks into syllables before the consonant.	If the last syllable of a word ends in le, the consonant preceding the le usually begins the last syllable.
		When the first vowel element in a word is followed by th, ch, or sh these symbols are not broken when the word is divided into syllables and may go with either the first or second syllable.
		In a word of more than 1 syllable, the letter v usually goes with the preceding vowel to form a syllable.
	When a single vowel falls between 2 consonants in a word or syllable, the vowel usually has the short sound.	When a word has only 1 vowel letter, the vowel sound is likely to be short.
		When there is one e in a word that ends in a consonant, the e usually has a short sound.
		When the last syllable is the sound r, it is unaccented.

When x appears at the beginning of a word, it makes the sound at the beginning of xylophone.

When gh appears at the beginning of a word, it makes the g sound heard in ghost.

When x appears at the end of a word, it has 2 sounds - k and s together

When mb occurs at the end of a word, the two letters usually have a single sound.

When lk occurs at the end of a word, the two letters have a single sound.

When y appears at the beginning of a word, it is a consonant with the same sound as the y in yellow and yet.

When g appears at the end of a word, it makes the sound as in tag.

An i followed by ld or ght makes a long sound.

An o followed by lk makes a long sound.

When a word ends with er, ir, ur or with those letters followed by other consonants, these letter pairs make the same sound as er in her.

	When y or r follows er, the vowel sound is similar to the one heard in very.
	When a vowel is followed by final re, the vowel does not have either the long, the short or the r controlled sound.
	When a word ends with or, or with or followed by one or more consonants, the or usually sounds like the or in torn.
	The vowel a has the same sound when it is followed by l, ll, or lk.
	One vowel at the end of a word usually has a long sound.
	When a vowel comes before nd or ld, it can have a long or short sound.
	When -ed, -ing, -er, or -le is added to a word with one vowel and one final consonant, the final consonant is doubled.
Unstressed vowels usually have the schwa sound.	The vowel sound in an unaccented syllable may be the schwa sound.
	Double consonants at the end of a word make 1 sound.
	If a word ends in ge, the sound of j is heard.
When words have 2 or more syllables and end with the letter y, the y usually stands for the long e sound.	When a word of 2 or more syllables ends in a y, the y usually stands for long e.

When you find 2 or more consonants between the vowels, try decoding the first vowel with its short sound first.

The letters qu spell the sound heard in quite.

The number of vowel soundsin a word represent the number of syllables found in that word.

A consonant cluster is 2 or more sounds that are blended together so that they almost sound like 1 sound when the word is pronounced.

Two vowel letters can stand for 1 vowel sound.

Two letters in the English language can represent 1 sound.

The 2 letters o and a together stand for a single vowel sound.

The 2 vowel letters ea usually stand for the sound of short e.

If one consonant falls between 2 vowels in a word, try a long vowel sound first. Then try short if that does not make sense.

If 2 or more consonants fall between 2 vowels in a word, divide the word between the 2 consonants to make 2 syllables. The first vowel is usually short.

The 2 vowel letters oo represent 1 sound. They may be found at the middle or end of a word.

One vowel sound can be represented by many spellings.

Digraphs consist of 2 letters that stand as 1 sound.

The c spelling and the k spelling for k may appear at the beginning, middle, or end of a word, while the ck may appear at the middle or end.

CONCLUSIONS

1. Phonics generalizations are included in current basal reading materials.
2. More phonics generalizations appear in current basals than were identified in those basals in a 1982 study.
3. Wording of the generalizations suggest that linguists have had an impact on phonics terminology in the elementary grades.
4. Rewording was noted in four of the 18 generalizations that Clymer found applicable at least 75 percent of the time. Many Clymer found less useful have been deleted.
5. Five generalizations are included in all three of the studies. Eleven of the generalizations identified in this study have not appeared in previous studies.

REFERENCES

Anderson, R. C., Hiebertl, E. H., Scott, J. A., and Wilkinson, I. A. G. (1985). Becoming a nation of readers: The report of the commission on reading. Washington, D.C.: National Institute of Education.

Clymer, T. (1963). The utility of phonic generalizations in the primary grades. The Reading Teacher, 16, 252-258.

Ingersoll, G. M., & Smith, C. B. (1984). 1983 report on trends in the use of educational materials: Traditionals. Dayton, OH: The Mazer Corp.

Sorenson, N. L. (1982). A study of the reliability of phonic generalizations in five primary-level basal reading programs. Unpublished doctoral dissertation, Arizona State University, Tempe.

EVALUATION OF TEACHING WITH DISCUSSION:
A TRIANGULATED INQUIRY

DONNA E. ALVERMANN
University of Georgia

The purpose of this study was to derive alternative interpretations of the data obtained from asking the question: What function does whole class discussion of assigned textbook readings serve?

Perspective

According to a recent survey of methods textbooks and journals (Gallagher & LaZansky, 1983), teachers generally hold fast to the assumption that discussion plays an important role in helping students comprehend what they read. As a topic of empirical investigation, however, teaching with discussion has received little attention since Taba's seminal work in that area during the early sixties. In fact, Tierney and Cunningham's (1984) review of the research on teaching reading comprehension since 1970 included only three studies directly related to classroom discussion.

That teachers continue to espouse the value of discussion despite very little confirming empirical evidence may be explained in part by the high regard they reportedly hold for two common instructional goals of discussion. Those goals, identified by Gallagher and LaZansky (1983) in their survey of the literature, are to:

(1) insure that all pertinent content information is made available to students

(2) heighten and extend students' interaction with ideas (p. 5).

The extent to which these two instructional goals guided one teacher's use of whole class discussion was the object of investigation in this study.

Methodology and Data Sources

Triangulation (Cicourel, 1974) was the methodology used to derive alternative interpretations of the function of whole class discussion. As an ethnomethodological approach toward understanding interpersonal events in the classroom, triangulated inquiry provided a means of evaluating from multiple perspectives one teacher's use of discussion. Compared to methods that employ a singular perspective, triangulated inquiry enables the researcher to explore whatever plausible interpretations are available from the differing participant perspectives. It also serves as a cross-validation check. Using this methodology, classroom discussion in one setting was studied from three stances of participant observation as described by Sevingy (1981): full participant, complete observer, and observer-as-participant.

Full participant. In this phase of data collection, group membership as a student in the first half of an 8-week summer course offered by the Sociology Department in the College of Arts and Sciences at the University of Georgia was adopted. Criteria for selecting the course included: the instructor's use of discussion on a regular basis and unfamiliarity with the instructor and the subject area. Since there was no wish to be treated differently from other students in the class, the research intention was not made known to the instructor or other members of the class until midway through the course. (The university human subjects committee had approved the project, however.) A notebook which was used for taking class notes conveniently provided a means for recording observation notes as well. The observation notes consisted of a log of the content of the discussion sessions, introspective analyses of the discussion process, samples of dialogue between the teacher and students, and descriptions of nonverbal behaviors exhibited during the discussion sessions.

Complete observer. After revealing the research interest in discussion during the second half of the 8-week course, audiotaping and key informant interviews were added as data sources. For two weeks these sources took the place of full participant observation notes.

Observer-as-participant. During the last two weeks of the course, the stance of observer-as-participant was assumed. In that role, interaction with both the instructor and students outside of the regularly scheduled class time was begun (e.g., during breaks, lunches, scheduled appointments, and walks to the library). Dual accounts of observations, accompanied by other participants' observations, formed working hypotheses to which two key informants' reactions were invited.

Data Analysis

Data analysis was divided into two phases: field (on going) and post-field (summative). In the field phase, discussion events were recorded in a journal, put on the word processor after class, coded as to emerging categories, and analyzed for whatever questions were generated that could guide subsequent observations during the next class period. In the post-field phase, field notes and audiotaped interview transcripts were duplicated, indexed, and cross filed. Data were analyzed using the constant comparison method of Glaser and Strauss (1967), a strategy for combining inductive category coding and simultaneous comparison of all events across all categories. As new relationships were discovered in this continuous refinement of categories, they were elaborated into an essay format and recorded in a separate journal.

Results

The function of whole class discussion changed as the 8-week summer course progressed. Initially, the teacher structured all discussions so that students felt free to interact with each other with little or no feedback from the teacher. Class participation was high, and the majority of the students kept up with the assigned readings and participated freely in the post-reading discussion sessions. After the midterm exam, however, attendance dropped dramatically. Of those students still in attendance, only two participated in discussions on a regular basis. Participant observation notes and transcripts of interviews with key informants revealed that students perceived a discrepancy in the teacher's reward system; i.e., divergent thinking was expected in class discussion, but convergent thinking in test taking. How this perceived discrepancy in the teacher's reward system was instrumental in altering whole class discussion patterns is described next.

Prior to the midterm exam, the teacher's encouragement of whole class discussion had served the two instructional goals described earlier by Gallagher and LaZansky (1983): i.e., it made available all pertinent content information, and it fostered student interaction with ideas. In fact, one of the discussion fostering categories that seemed to emerge from using the constant comparison method of Glaser and Strauss (1967) was the Sustaining Discussion category. A property of that category was critical questioning, which served as the means for stimulating discussants to react to, judge, and evaluate ideas presented by the text and by others. An example of how the teacher sustained discussion using critical questioning is presented here:

T. Can anyone infer from the reading how Max Weber viewed the control of bureaucracy?

Eric. Wait. First, I see a problem with Weber's three forms of authority -- traditional, charismatic, and rational.

T. Interesting observation. Anyone care to comment on Eric's statement?

Jolyn. Well, okay, all three can become, uh, routinized. (Looks toward Debbie, as if for support.)

Debbie. It seems like a charismatic leader is the one who always acts as a checks and balance; for example, the leader, um, Jim Jones made people stop and take a look at cults in the U.S.

T. (moving over to stand by Debbie's desk) How can you relate your example to Jolyn's notion of the routinization of Weber's three forms of authority?

The discussion between Debbie and Jolyn continued for more than three minutes. When the teacher did interrupt, it was not to give his opinion of the "rightness" or "wrongness" of the students' comments, but rather to introduce the next reading assignment, which he said provided information that could extend the discussion even further.

In contrast to this earlier exchange between teacher and students is the next segment of a dialogue, which was taped two weeks after the midterm grades had been posted. This time the teacher asked questions that required only a surface level processing of the information in the assigned article. The students, in turn, provided little more than the obligatory responses.

T. Okay, anyone want to make any comments about Perrow's article? (in reference to their assigned reading of "Normal Accident at Three Mile Island.")

T. (Throwing a piece of chalk in the air, catching it, and walking around his desk several times) Patti, what do you think? Did you find the language straight forward? (in reference to Perrow's frightening account of the potential for more nuclear accidents like Three Mile Island.)

Patti. Yes.

T. Did the topic cause you to lose any sleep?

Patti. No.

T. What's the point of reading this article anyway, Deborah?

Deborah. So we don't forget the danger of nuclear power.

T. I can't argue with that.

T. But Perrow is saying more than that.

T. (waits for a student to respond, but no one does) Why is the normality he's talking about predictable?

Patti. Accidents do happen.

The teacher made a few more general statements about the Perrow article and then said: "We're finished with Perrow. I don't think there's much more we can say. The complexity of the system results in people not being able to cope with it."

Impression of changes validated. The impression that the function of the whole class discussion approach to learning had changed after the midterm exam was validated from multiple perspectives. A comparison of the field notes from the first four weeks as full participant (prior to the exam) with field notes from the last half of the course revealed a decided change in function. Prior to the exam, the teacher would frequently praise students for carrying on a discussion in a manner that showed they were the "sphere of work" and not the least bit dependent on him as "a delivery service of knowledge" (from field notes 6-19, p. 10, lines 14-16). Or, he would listen intently for relatively long periods of time as students debated back and forth about issues that had been raised (from field notes 7-3, p. 5, lines 21-25). Whole class discussion functioned as a strategy for stimulating thinking, challenging ideas of others as well as ones own. Students were free to agree or disagree, and almost everyone had an opinion.

After the midterm exam, however, neither the teacher nor students appeared capable of sustaining a discussion long enough to examine a set of ideas. Attempts at even initiating a discussion often met with failure. Consider, for example, the following attempt (from field notes 7-24, p. 3, lines 3-6):

T. Any questions or comments on anything in the Chapter?

T. (after several seconds of complete silence) On all of it?

T. (again, after an uncomfortable silence) I don't know whether it's a problem of people being shy or whether it's the density of the material.

Key informant interview data suggested that the problem in initiating discussions could not be attributed to either the shyness of the individuals or the density of the material. Rather, as Dan (interview #14, p. 3, lines 13-18) pointed out, "I don't talk because I don't like how the class is going. Talking just to be talking, like answering his (the teacher's) questions, isn't where my head's at. Used to be our opinions and ideas were fine in class, but come exam time, and no -- well I should have known it would end up that way." Kristi (interview #3, p. 1, lines 9-10), voiced a similar appraisal of the situation: "If we're gonna get graded on the book -- what it says -- fine, but then why waste the time we spent arguing points -- our ideas, you know." Cindy (interview #9, p. 6, lines 8-9) was the least analytical in describing how the function of whole class discussion had changed. She merely noted that "Without planned discussion sessions, this course is like all the other courses I take. No one talks or gets involved, really." When it was mentioned that the teacher continued to have discussion sessions each day, frequently as class openers, Cindy looked puzzled and said "But not really." To her way of thinking, a planned discussion session made explicit the involvement of students interacting with issues that they had identified from the reading, whereas discussion per se might simply mean the teacher discussed (lectured on) the assignment or asked specific questions of specific students.

All three key informants mentioned a change in the discussion patterns following the midterm. Dan (interview #14, p. 1, lines 14-18) attributed the change to the teacher's need to rush through the course content. He thought that the reluctance on the part of the students to participate was only partially due to their realization that convergent thinking, not divergent thinking, was rewarded on the midterm. Kristi and Cindy, on the other hand, viewed the decrease in student participation during whole class discussion as an indication of the students' collective rejection of the notion that discussion could function as a learning strategy that paid off at test time. Kristi (interview #3, p. 1, lines 7-8) said that she now preferred the type of discussion signaled by the teacher's question, "Do you have anything in the readings that you want me to clarify?" She thought that she would be able to make better use of that type of information on the final. Cindy (interview #9, p. 1, lines 11-13) admitted that she had come to feel more secure when the teacher announced, "What I want to give you is this list of three distinctions that separate simple, bureaucratic, and technical control" (whereupon everyone in the class picked up a pencil and began to write).

Teacher's impression of the change. From the teacher's perspective, whole class discussion was never meant to function as more than a means for motivating students to read their assignments (interview #2, p. 2, lines 3-4). During a scheduled office appointment with the teacher toward the end of the seventh week, I asked him to react to my hypotheses that the results of the midterm exam served to turn students away from active participation in whole class discussion. He seemed genuinely surprised by this interpretation of lowered class participation and attributed the cause instead to students' falling behind in their readings (interview #15, p. 2, lines 17-19). He also said that he felt a certain pressure to get through the material at a bit faster rate following the midterm (line 23). This latter observation supported Dan's earlier remark.

Although the teacher rejected the notion that the function of whole class discussion changed when the students no longer perceived their participation as benefitting them on the midterm, he did indicate that he sensed he had to "work harder to pull fewer and fewer opinions from the students" (interview #15, p. 3, line 19). As before, however, he tied this observation to the students' general lack of interest in reading the assigned chapters.

Conclusions

Evaluating teacher use of discussion through a triangulated inquiry method permitted a cross-check of preceptions held by students and teacher of the function served by whole class discussion of assigned textbook readings. Although students and the teacher apparently were not in agreement regarding its function from the start, whole class discussion (at least initially) served the two common instructional goals of making available pertinent content information and fostering student interaction with ideas. Although the perceived reasons for the change in the function of whole class discussion following the midterm exam varied, there was general agreement that a change nonetheless had occurred and that the two instructional goals of discussion were no longer fulfilled.

Finally, through the cross-checking of systematic observation notes and interview data, the complex manner in which discussion functioned began to show itself. Most of the students, prior to their first exam, viewed whole class discussion as a means for examining textbook issues that concerned them. After their first exam, they viewed discussion as a means for learning what the teacher deemed important from the text. The emphasis moved toward convergent thinking, not divergent thinking. The teacher, meanwhile, never intended for whole class discussion to function in a capacity other than that of motivating students to read their assignments. Through a stripping away of the discussion's outside coverings, however, it was possible to look at alternative interpretations of the same phenomenon.

REFERENCES

Cicourel, A. (1974). Cognitive sociology: Language and meaning in social interaction. New York: Free Press.

Gallagher, M. & LaZansky, J. (1983). The effects of small group discussion on student acquisition of content information. Unpublished paper, Center for the Study of Reading, University of Illinois.

Glaser, B. B., & Strauss, A. L. (1967). The discovery of grounded theory. New York: Aldine.

Sevigny, M. J. (1981). Triangulated inquiry -- A methodology for the analysis of classroom interaction. In J. L. Green & C. Wallat (Eds.), Ethnography and language in educational settings (pp. 65 - 85). Norwood, NJ: Ablex.

Tierney, R. J., & Cunningham, J. W. (1984). Research on teaching reading comprehension. In P. D. Pearson (Ed.), Handbook of reading research (pp. 609-655). New York: Longman.

THE ROLE OF SELECTED PRINCIPALS IN THE EVALUATION OF THE TEACHING OF READING

REBECCA M. MOORE

SALLY M. BOYETT

SYLVIA MCCOY CARTER
University of Georgia

A recurring theme found in school effectiveness literature is that principals in effective schools are strong instructional leaders and therefore, play a crucial role in influencing the school's high achievement (Wellisch, et al, 1978; Rutter, et al, 1979; Edmonds, 1979; and reviews: Shoemaker and Fraser, 1981; Purkey and Smith, 1982). This finding is mirrored in recent studies of effective principals (Blumberg & Greenfield, 1980; Rutherford, 1985; and reviews: Sweeney, 1982; DeBevoise, 1984).

In the literature on effective principals, the concept of instructional leadership encompasses those actions or behaviors of a principal which promote growth in student learning. The most recent investigation of effective leaders (Rutherford, 1985) supports the findings of previous research. For five years a research team studied the leadership skills of elementary and secondary principals in Texas. Rutherford discussed the distinctions between more effective and less effective principals in relation to five essential qualities of effective principals. These essential qualities have been clearly and consistently identified by earlier research. Effective principals 1) have clear, informed visions of what they want their schools to become- visions that focus on students and their needs; 2) translate their visions into goals for their schools and expectations for the teachers, students, and administrators; 3) establish school climates that support progress toward these goals and expectations; 4) continuously monitor progress; and 5) intervene in a supportive or corrective manner, when this seems necessary.

School effectiveness research indicates clearly that a significant factor in the school's instructional success is the principal's involvement. Although limited in number, investigations of effective/quality reading programs have reaffirmed strong administrative leadership in those schools with high levels of reading achievement (Rauch, 1974; Guthrie, 1977; McNinch & Richmond, 1977; Manning & Manning, 1981). "The key to the improvement of reading rests with the principal. By the very nature of the position, the principal is responsible for providing the impetus to improve the school reading program" (Bernard & Hetzel, 1976, p. 386).

According to Egherman (1981), the following leadership qualities have been shown to make a difference in increasing reading skills:

a) In order to know what to look for when observing classroom performance and how to assist teachers, principals must have experience in reading instruction or an understanding of the process.

b) The principal must commit more time observing reading instruction and providing assistance to teachers.

c) The principal must communicate to the staff that reading is an important instructional goal.

In conclusion, the following summary statements are made in reviewing the literature relating to administrators as effective instructional leaders in the school reading program:

1. A common element in effective schools is strong instructional leadership on the part of the principal or another staff member.

2. As a strong instructional leader, the principal plays a crucial role in influencing the school's high achievement.

3. The principal is responsible for providing the impetus to improve reading in the school.

4. Effective reading programs in schools with high levels of reading achievement have strong leadership from the principal.

5. The principal is responsible for a consistent, ongoing system of program and staff evaluation.

Procedures

Using the 1984 Georgia Public Education Directory 159 elementary and middle school principals were arbitrarily selected for study from metro Atlanta, Columbus, Gainesville, Valdosta, and the five counties surrounding Athens. An attempt was made to regionally accommodate the selection of these public school principals from some of the larger population centers in Georgia.

One hundred fifteen elementary and middle school principals returned the questionnaire, with 48 principals agreeing to be contacted for further information. Of the 48 principals contacted with a follow up questionnaire, 29 of them responded with additional information.

Based upon what research shows to be the fundamental characteristics of an effective instructional leader, a one-page, multiple-choice, brief answer questionnaire was mailed with a self-addressed stamped envelope to 159 public school principals throughout the state of Georgia. Designed to be brief and nonthreatening, the questionnaire attempted to determine the principal's years and nature of classroom teaching experience, the number and variety of reading education coursework taken, and the procedures used to evaluate the teaching of reading. An opportunity was given for each principal to agree to be contacted for further information.

A second multiple-choice questionnaire was then mailed with a self-addressed stamped enveloped to the 48 principals who agreed to be contacted for further information. They were requested to respond to the perceived quality of their school's reading program and library, as well as their involvement with reading textbook selection and evaluation. In addition, these principals were given the opportunity to describe their involvement with staff development in the reading program, parental programs that support reading and the school programs that emphasize reading.

Not purporting to be survey specialists, the researchers recognized several limitations with both survey instruments. As with any self-reporting instrument, the respondents' perceptions may not always be an accurate description of reality.

The questionnaires lacked specificity in the areas of years in current assignment as principal, discrimination between inservice and college level reading coursework, and the co-ordination of the nature of teaching experiences with current assignment. Despite these limitations, 72% of the first questionnaire were returned with 62% of the second follow-up questionnaire returned.

Results and Discussion

Having computed percentages of item responses, three main areas of concern were examined: classroom teaching experience, reading education coursework, and evaluation of the teaching of reading.

Of the 115 respondents, 66% of them had five years or less elementary classroom teaching experience, 89% had five years or less middle school classroom teaching experience, and 84% had five years or less secondary classroom teaching experience. While not able to explicitly determine which principals had which kind of classroom teaching experience, in general, the majority of the principals had five years or less experience in the classroom. The number of respondents/principals with no classroom teaching experience in one or another of the three specified levels is in contrast to traditionally expected experiential background.

Of the 115 principals 67% have had 10 years or less experience as a principal. For the most part, the principals surveyed in Georgia tend not to fit the traditional profile of an elementary principal: middle-aged or older and well-experienced. In contrast, the majority of the principals surveyed are relatively young with limited teaching experience.

Several implications may be drawn from the findings with regard to classroom teaching experience of surveyed principals. The principal as the instructional leader may be called upon to provide staff development for the reading program. Due to a lack of classroom teaching experience, the school system's central office may find that such young principals are in need of training in order to better meet the staff development needs of their school's reading program. As Becoming a Nation of Readers recommends both the novice and experienced teacher have needs that must be met in order to encourage reading in their classrooms.

Research has shown that the principal provides the impetus to programs that improve reading. In order to provide that impetus principals need to be academically prepared, as well as experienced in classroom teaching. This study's findings show that 28% of those principals responding have had one or less inservice/ graduate level courses in reading education. While the conclusion may appear to be that 28% of these principals could not be certified to teach in Georgia's elementary classrooms, in reality, these respondents may have been middle school principals who need but one reading course to teach in the middle school classroom. No reading education courses are required to obtain a leadership certificate in the state of Georgia.

As with a principal inexperienced in the classroom, a principal not adequately prepared academically may not be as strong in providing staff development as one more familiar with teaching of reading. There is a genuine need to encourage principals to pursue valid reading education courses that will better prepare the instructional leader to administer and implement the reading program. In order to compensate for a lack of direct knowledge, an effective principal must be aware of those resource people upon whom he may call for assistance in building and maintaining effective reading programs.

When asked how reading programs were evaluated in their schools, 97% of the principals returning the questionnaire reported they used direct observation of the teaching lesson. Other methods of evaluation employed included checklists of teacher behavior and program characteristics, individual conferences, teacher self-evaluations and student evaluations. Recognizing the limitations of a self-report instrument, judgment is not being made concerning the techniques reported.

The evaluation procedure options available on the questionnaire were not intended to be mutually exclusive of each other. Conceivably, a principal may employ several techniques during the evaluation of the teaching of reading. Each principal chose the evaluation procedures based on his or her interpretation of the evaluation tools listed. While a variety of evaluation instruments were discussed, only a small percentage of respondents chose to send back a sample of their evaluation instrument. Of those sent, the evaluation tool tended to be a standard form used within a school district and not specifically related to the teaching of reading. From the follow up questionnaire sent to 48 principals, 62% of the 29 respondents rated their reading programs as "above average." Understandably, a self-report instrument measures perceptions and does not give the basis upon which their decisions were made. Common sense would lead one to assume only those principals with above average reading programs would report their findings as such.

Research points to the importance of employing a consistent ongoing system of evaluation in schools that are striving to become more effective places of learning. Principals need to be prepared in the selection and utilization of evaluation procedures for the teaching of reading. Classroom teaching experience and academic preparation are two key factors in evaluating the reading program. A limitation in either area breaks down the evaluation process. Staff development for the principal may be conducted to insure the principals have a stronger basis upon which to evaluate the teaching of reading in their elementary and middle schools.

References

Anderson, R.C., Hiebert, E.H., Scott, J.A., & Wilkinson, I.A.G. (1984). Becoming a Nation of Readers: The Report of the Commission on Reading. Washington, D.C.: NIE.

Barnard, D. P., & Hetzel, R. W. (1976). The principal's role in reading instruction. The Reading Teacher, 29(4), 386-388.

Blumberg, A., & Greenfield, W. (1980). The effective principal: Perspectives on school leadership. Boston: Allyn and Bacon.

De Bevoise, W. (1984). Synthesis of research on the principal as instructional leader. Educational Leadership, 41, 15-20.

Edmonds, R. (1979). Effective schools for the urban poor. Educational Leadership, 37, 15-27.

Egherman, D.A. (1981). A short course in improving reading skills. NASSP Bulletin, 65, 75-81.

Georgia Public Educational Directory (1984). Atlanta: Department of Education.

Guthrie, J. J. (1977). Research views: Recreating successful reading programs. The Reading Teacher, 30, 952-955.

Manning, G. L., & Manning, M. (1981). What is the role of the principal in an excellent reading program? Principals give their views. Reading World, 21, (2) 130-133.

McNinch, G. H., & Richmond, M. G. (1977). Teachers' perceptions of their principal's role in elementary school reading instruction. Reading World, 17, 55-63.

Purkey, S. C., & Smith, M. S. (1982). Too soon to cheer? Synthesis of research on effective schools. Educational Leadership, 40, 64-69.

Rauch, S. J. (1974). Administrator's guidelines for more effective reading programs. Journal of Reading, 17, (4), 297-300.

Rutherford, W. L. (1985). School principals as effective leaders. Phi Delta Kappan, 66 (13), 31-34.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). Fifteen thousand hours. Cambridge: Harvard University Press.

Shoemaker, J., & Fraser, H. W. (1981). What principals can do: Some implications from studies of effective schooling. Phi Delta Kappan, 63 (3), 178-182.

Sweeney, J. (1982). Research synthesis on effective school leadership. Educational Leadership, 39, 346-352.

Wellisch, J. B., MacQueen, A. H., Carriere, R. A., & Duck, G. A. (1978). School management and organization in successful schools. Sociology of Education, 51, 211-226.

NEW WAYS TO USE STUDENT WRITING SAMPLES IN READING DIAGNOSIS

PATRICIA H. DUNCAN
Virginia Commonwealth University

This paper introduces a rationale for using writing samples in reading diagnosis. It is predicted on an understanding that effective diagnostic practices, whether they be in classroom or clinic, relate reading behavior to other aspects of learning.

Regular components of a diagnostic session typically include testing in areas of oral language, mathematical computation and areas of perception. When writing is included as part of the assessment, it generally is administered as a small segment of a comprehensive battery for individual reading diagnosis. Samples often are no more than sentences and the evaluation focuses on spelling miscues or indices of syntactic maturity.

Writing samples are considered very useful in assessing total language competence. Given the appropriate orientation, writing products can provide meaningful insights into a variety of cognitive behaviors. This paper will review cognitive and linguistic competence revealed by a piece of writing. Formats for generating a good writing sample during the diagnostic session are suggested. Finally, the paper recommends assessment techniques gleaned from research on this topic which might be incorporated into diagnosis for understanding of reading status.

Using Writing Samples in Reading Diagnosis.

While recognizing that reader and writer differ in the manipulation of discourse (Moffett, 1983), diagnosticians are able to gain a broader perspective of a reader's language competence by looking at his/her writing product. According to Moffett (1983), the reader restructures consciousness by a process of mediating the author's thoughts. In composition the writer transforms thought on several levels by transcribing inner speech, self-dictation. This part of composition shows what can be done with inner speech through focusing, revising, and editing. When the writer selects ideas or images and presents them in coherent continuous text, he/she reveals the capacity for mediation. (p. 320-321).

Because writers generally read and re-read during writing it is possible to infer from the product the various levels of language competence. The writer reveals his/her available schemata, capacity for organization, syntactic sophistication, knowledge of spelling patterns, and control of punctuation. Tone and style often give clues to personality and maturity.

Features that make writing useful as a component of expanded reading diagnosis can be explained through several arguments:

1. During reading testing the client is called upon to perform chiefly in the abstract modes of speaking, reading, and listening. Writing, while it is more complex, does allow for tangible operation with pencil and paper.
2. The client is given the opportunity to perform in a mode which can yield strong ego-gratification. The success of language experience approaches to reading is testimony to the importance of seeing one's ideas in print. When demands for mechanical perfection are removed, a similarly reinforcing situation is available as the client does the writing.
3. Society has not determined so definitely what constitutes success in writing; therefore, a client who has a history of failure as a reader may welcome the change to the expressive written mode. Acceptance of his/her composition content may lead eventually to greater interest in form and convention. This has been demonstrated many times through tutorials in which the lessons are a combination of reading and writing instruction.
4. The role of the examiner administering tests changes during the writing task from direction giver to audience. In some instances the examiner may enter into a peer relationship in which he/she also writes and shares.

Two Models For Analyzing Writing

As with reading evaluation, writing evaluation is influenced by the nature and purpose of the task. The form of writing requested (dramatic, reporting, research, personal, technical, etc.), the developmental expectations and the relevance to the real world are all aspects to be considered in analyzing a writing sample. The mode chosen for writing-expressive, transactional or poetic (Britton, 1982) dictates the appropriate diagnostic perspective. In the expressive or personal mode a student may write actively with less attention to form. On the other hand, beginning attempts at poetic writing may offer stilted content while the student tries to master form.

When purpose for writing and choice of mode of expression have been decided, the diagnostician then can analyze the writing from a selected perspective. A useful model cited by Cooper (1973) was developed for holistic evaluation by the London Association for the Teaching of English in 1965. The main criteria of this model as paraphrased by Cooper are:

1. Realization: the extent to which the writing directly reflects the writer's own experience (sincerity, spontaneity, vividness).

2. Comprehension: the extent to which a piece of writing shows an awareness of audience and can thereby be understood, permitting some consensus of response.

3. Organization: the extent to which a piece of writing has shape or coherence.

4. Density of Information: the amount of unique and significant detail.

5. Control of Written Language: the extent of control over the special forms and patterns of written syntax and rhetoric.

Using this model the writer is judged first for creativity and ability to use internal schemata to convey thoughts. In criteria 2 and 3 the writer is evaluated for higher order comprehension abilities. He/she is asked to take the role of reader (application) and to organize the product into a coherent unit (analysis/synthesis). In the fourth criterion the writer is judged for amount and flexible use of vocabulary. Degree of language expansion also comes into play in this criterion. The last criterion, 5, is an indicator of word knowledge and grammatical usage. Spelling patterns, sentence structure and punctuation are thus observed after content, creativity, and comprehension. This model gives the writer the benefit of having his/her ideas examined before looking at mastery of the surface features of language. A sixth category could be added to the London Model: Control of visual motor behaviors: the extent to which the writer demonstrates developmentally adequate handwriting (letter formation, spacing, proportion, slant, etc.).

Another British writing assessment model, developed through the Crediton Project, presents a guide to developmental analysis of writing along four task dimensions. Wilkinson, Barnsley, Hanna, and Swan (1983) summarized their system of analysis as follows:

> Cognitive: A movement from an undifferentiated world organized by mind, from a world of instances to a world related by generalities and abstractions.
>
> Affective: Development is seen as being in three movements - one towards greater awareness of self, a second towards greater awareness of neighbor as self, a third towards an interengagement of reality and imagination.
>
> Moral. "Anomy" or lawlessness gives way to "socionomy" or rule by a sense of reciprocity with others, which finally leads to the emergence of "autonomy" or self-rule.
>
> Stylistic: Development is seen as choices in relation to a norm of the simple literal affirmative sentence. Features, such as structure, cohesion, verbal competence, syntax, reader awareness, and appropriateness are studied. (p. 873)

Of course, this model was not intended for diagnosis but as a device for guiding teaching in the school districts in the United Kingdom, Australia and Canada. As did the London model, the Crediton model places observation in an appropriate hierarchy with writing conventions viewed as only one component of the writing product. This model brings affect and values into focus along with the cognitive features of composition. Sometimes feelings, attitudes and values are overlooked in reading diagnosis; yet, they may play an important part in motivation to read and in the choice of material to be read. A sample designed to capture these features as well as content and conventions of written language would be helpful indeed.

Generating Writing Samples During Reading Diagnostic Sessions.

The ideas presented here are examples of activities which can be implemented in one-to-one situations or as group activities. Some are more appropriate for young children; others are geared for older students or adults. All of the tasks described require some form of prewriting discussion, time for drafting and a short sharing experience following the writing. Whenever possible, the examiner could share in the activity along with the student at least for a short period.

This lends credibility to the tasks and tends to reduce tension as the writer becomes less conscious of the examiner. The examiner gives brief time limits for the writing and clarifies strategies for coping with hard-to-spell words. The examiner remains unobtrusive as the writing progresses, quietly maintaining attention to task.

Narrative/Descriptive Tasks Include:

1. Picture Stimulus. Select a picture which will relate to the client's interests--one that has the potential for a general detailed description of a story. Young children seem to enjoy the works of Anno, Scarry and Spier which are full of detailed activity. Older clients may relate to a scene or a picture of a famous person. Prewriting discussion might start by using one or more of the following:

   --Tell me about this picture
   --What do you see in the picture?
   --Does the picture remind you of anything that ever happened to you?
   --What if the _____ in the picture were _____?

   Then the client is asked to generate a writing sample.

2. Object Stimulus. Assemble a collection of toys or objects. Ask the client to choose one to write about. Prewriting discussion focuses on identifying the object, telling who might use it or be associated with it and what event or problems might occur related to the object. Give a five minute writing period. Share product after the writing.

3. Story Starters. Use a group of pictures, objects, or toys. Have the client select one from the group. Read the situation story starter aloud to the student. Ask what he/she thinks will happen next. Ask the client to write for five to ten minutes. Share the writing.

4. Personal Narrative. Draw a question from your interview with the student or use a standard human theme, such as, "Have you ever wanted to do something but were afraid to do it?" Another useful question is "Do you remember a time when you were little and you got into trouble?" Talk about the incident briefly, and then ask the client to write for five to ten minutes. Ask clients if he/she would like to share the story with you.

For less able and younger writers:

5. Draw and Write. Have the client draw a picture related to an item of interest or a recent experience. Have the client draw for a few minutes, then ask the client to write about the topic for about five minutes. Indicate that the picture may be finished later. Keep the picture for evidence of pictographic or related ideas.

6. Story Skeleton. For the young child who is reluctant to write a simple story skeleton may help. A common format is described by Tiedt and Tiedt (1978, p. 2):

Who Goes There?

Complete each sentence to make a story.

1. After school I went ______________________ ______________________.
2. Suddenly I heard ______________________ ______________________.
3. Behind a tree I saw ______________________ ______________________.
4. I felt happy because ______________________ ______________________.

Additional procedures can employ Reflective/Analytic Writing:

7. Personal Writing. This requires freedom and latitude on the part of the observer. Prewriting discussion might begin with questions such as:

   --Can you think of one of the best days you have ever had?
   --What is your favorite possession?
   --Can you think of an accident that happened to you or someone close to you?

   Then instruct client to reflect on the topic and write for a specified time.

8. Transactional Writing. Ask the client to write about a specific topic. Some examples are:

   --You told me you can make corn bread. Will you write down the directions for me?
   --Write about what it was like to ride in the dune buggy this summer.
   --(Draw an incomplete shape S) Make this into a picture and write what it is.

This type of task asks the writer to provide information. It can be helpful in assessing sequential organization, logic, and language clarity.

9. Reciprocal Writing. This interesting task borrows from the work of Peter Elbow (1973) and Mary Kitagawa (1982). It sets up a writing situation in which client and examiner write for a short time about an agreed-upon topic. Then, they exchange papers and respond in writing to the work of the other. The exchange takes place a second time. This task gives the client immediate support and feedback and demonstrates the links between reading and writing.

Finally, use Writing From A Literature Stimulus:

10. Read and Write. The examiner reads a short stimulating segment to the client and then, depending on the nature of the story, asks the client:

    --What might happen to the protagonist in the future?
    --What might have been another ending to the event?
    --Imagine yourself in the story. What would you have done in this episode?

This stimulus gives structure and direction to older writers. A series of topic sentences may also be developed to guide the writing.

Overview of Assessment Techniques for Writing that Have Implications for Reading Diagnosis.

Knowledge of the types of writing evaluation available helps the reading diagnostician to choose appropriate methods for analyzing writing samples. The National Writing Project and the National Assessment of Educational Progress have had profound influences on changes in the direction of writing evaluation. Atomistic evaluation, which counts discourse features, such as T-units, has been replaced by holistic scoring methods. Holistic scoring assumes that error counts alone cannot reflect language competency. In holistic scoring the product is rated as a whole against some prototypes or anchors which have been previously analyzed by using carefully defined criteria (Myers, 1980). Examples of holistic scoring as identified by Cooper (1977) are essay scale, analytic scale, dichotomous scale, feature analysis, primary trait scoring, general impression marking, and "center of gravity" response.

The value of these types of scoring is that they do not merely count features. They do require the evaluator to predetermine criteria for evaluation. The feature analysis and dichotomous scales seem particularly suited for inclusion in a diagnostic situation with time constraints. These types of scales allow the examiner to look at a piece of writing systematically. A simple scale for dichotomous scoring is suggested as a modified version of one developed by Cohen (1973):

Dichotomous Scale For Evaluating A Writing Sample

	YES	NO		
Content	___	___	1.	Ideas reflect personal experience.
	___	___	2.	Ideas are creative or original.
	___	___	3.	The writing follows a logical premise.
	___	___	4.	Ideas are expressed clearly.
Organization	___	___	5.	Topic or theme is followed throughout the writing.
	___	___	6.	Topic is adequately developed.
	___	___	7.	Details are relevant and well-ordered.
Mechanics	___	___	8.	Sentences flow smoothly.
	___	___	9.	Word choice is correct.
	___	___	10.	Grammar meets educational standard.
	___	___	11.	Spelling is developmentally adequate.
	___	___	12.	Punctuation errors are not excessive.

Another tool for evaluating writing is the post-writing interview. This format is useful in determining the client's knowledge of rhetoric in narrative writing. An interview format for use with first grade writers was developed by Duncan (1985) to evaluate children's concept of story in their own writing. The interview guide follows story grammar format and elicits the writer's knowledge of the elements in a simple episode.

Writing Evaluation
INTERVIEW FORMAT

Patricia Duncan, 1985

1. (Overview) TELL ME ABOUT YOUR STORY.
2. WHAT MADE YOU (Why Did You Decide To) WRITE ABOUT _____?
3. (Beginning Element) WHAT DID YOU THINK OF AS YOU BEGAN TO WRITE YOUR STORY?
4. (Formal Opening) WHY DID YOU DECIDE TO START YOUR STORY WITH ______?
5. (Characters) WHY DID YOU PUT ___________ IN YOUR STORY?
6. (Voice) ARE YOU IN THE STORY?
7. (Character Development) READ ME THE PART THAT BEST DESCRIBES _____? WHY DID YOU DECIDE TO WRITE THAT?
8. (Plot/Problem) WHAT DO YOU LIKE BEST ABOUT YOUR STORY?
   OR
   READ ME THE MOST EXCITING PART OF YOUR STORY. WHY DID YOU DECIDE TO WRITE ABOUT THAT?
9. (Ending) READ ME THE END OF YOUR STORY.
   OR
   WHY DID YOU END THE STORY THE WAY YOU DID?
10. (Audience) WHO WOULD LIKE TO READ YOUR STORY?

By using this interview process, one obtains clues as to the students' understanding of such abstract concepts as plot, theme, and character development.

Summary.

Writing samples can offer an expanded look at language competency. By comparing reading and writing behaviors, diagnosticians could gain a more comprehensive picture of the client's reading, ideational capacity, linguistic competence, and mastery of the mechanical conventions of spelling and punctuation. The examiner may also obtain clues for using reading and writing as modes to reinforce learning.

REFERENCES

Britton, J. (1982). Prospect and Retrospect: Selected Essays of James Britton. Montclair, NJ: Boynton/Cook.

Cohen, A. (1973). Assessing college students' ability to write compositions. Research in the Teaching of English 7, 356-71.

Cooper, C. (1977). Holistic evaluation of writing. In C. Cooper and L. Odell (Eds.), Evaluating Writing. Urbana, Illinois: National Council of Teachers of English.

Duncan, P. (1985). What do children understand about their own compositions? Paper presented at the Spring conference of the National Council of Teachers of English, Houston, TX.

Elbow, P. (1973). Writing Without Teachers. New York: Oxford University Press.

Kitagawa, M. (1982). Expressive writing in Japanese elementary schools. Language Arts, 59, 18-22.

Moffett, J. (1983). Reading and writing as mediation. Language Arts, 60, 315-322.

Myers, M. (1980). A Procedure for Writing Assessment and Holistic Scoring. Urbana, Illinois: National Council of Teachers of English.

Tiedt, S. & Tiedt, I. (1978). Language Arts Activities For the Classroom. Boston: Allyn and Bacon.

Wilkinson, A., Barnsley, G., Hanna, P. & Swan, M. (1983). More comprehensive assessment of writing development. Language Arts, 60, 871-881.

INDEX

Alvarez, John, 222
Alvarez, Marino, iii, 251
Alvermann, Donna E., iii, iv, 329
American Reading Forum, ii
Askov, Eunice N., iii, 74

Benton, Sidney E., 82
Blair, Timothy, 92
Bloomer, Richard, 286
Boarks, Nancy, iii, 306
Boyett, Sally M., 337
Bullock, Terry, 126

Calloway, Byron, iii
Camperell, Kay, iv, 30
Carlson, Deborah L., 44
Carter, Sylvia McCoy, 337
Casale, Ula, iii
Casbergue, Renee M., 194
Cassady, Judith, 59
Cloer, Carl Thomas Jr., iii, 212
Collins, Martha D., iii, 317
Czajkowski, Ania, 265

Dahlberg, Lucy Ann, iv
Dana, Carol, 41
DeSanti, Roger J., iii, 194
Dinnan, James A., 68
Duffy, Gerald, iv
Duncan, Patricia, 344

Early, Julie, 306
Elliott, Joan B., iv, 174
Erickson, Lawrence, 59
Evans, Deborah, iii

Flippo, Rona F., iv, 229, 296

Gentile, Lance M., 99

Hall, Ann, iii
Handler, Janet, 44
Harshbarger, Mary, i, iv, 149
Hays, Bernie, iii, 35
Henk, William A., 186
Herrmann, Beth Ann, iii, 233

Jennings, Robert, iii, iv
Jerrolds, Bob W., iii, 82

Kolker, Brenda, iii
Kossack, Sharon, iii, iv, 1, 222,278
Kurth, Ruth, iv, 38, 53

Laffey, James, 17
Laine, Chester, iii, iv
Laudermilk, Virginia, 278
Levesque, Elizabeth, 286
Lofton, Glenda, iii
Long, Judith L., 82
Lumpkin, Donavon D., i, iv, 167

Maclay, Connie, 297
Madden, Deborah Rosenerans, 126
Mallery, Anne, iii, 113
Manzo, Anthony, iii, iv
Martin, Michael, 286
Mason, George, iii, iv
McNinch, George, iv
Moore, Rebecca M., 337

Neville, Donald D., 240
Nolander, Kay, 265
Norton, Donna, iv

Otto, Wayne, iii, iv, 28

Paolitto, Anthony, 265

Peters, Charles, iii
Pulling, L. Vern, 68

Ransom, Peggy E., i, iv, 143
Raygor, Betty, iii
Rebottini, Sondra L., 132
Richmond, Katherine, 240
Risko, Victoria, 105
Rogers, Deborah J., 59

Sable, Eileen, 306
Searls, Evelyn F., iv, 240
Shaver, Judy C., 253
Smith, Patricia K., 132
Spache, Evelyn, iii
Stahl, Norman, iii, iv
Stice, Carole K., 98, 204, 305

Telfer, Richard, 205
Terwilliger, Paul, iii
Tomlinson, Louise M., 192, 283
Tonjes, Marian, iv
Townsend, Brenda S., 261

Wachter, Louise, 74
Wallenberg, John, 111
Weidler, Sarah D., iii, 297
Williams, Joan, iv, 156
Wise, Beth, 253
Wise, Kathleen Gaare -59

Yarbough, Betty, iii
Young, June Thomas, 92